

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:06:08 ; Search time 40 seconds
(without alignments)
940.272 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDGALPEINSARMYAGP.....SGVLVPPRPYVMPHSPAAG 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1949	100.0	391	2 B70608	probable PPE prote
2	1656.5	85.0	396	2 H70741	probable PPE prote
3	1583	81.2	393	2 C70568	probable PPE prote
4	779.5	40.0	393	2 G70929	probable PPE prote
5	751	38.5	409	2 A70932	probable PPE prote
6	737.5	37.8	403	2 H70931	probable PPE prote
7	736	37.8	423	2 B70931	probable PPE prote
8	731	37.5	421	2 H87056	PPE-family protein
9	705	36.2	408	2 G70925	probable PPE prote
10	702	36.0	391	2 B70625	probable PPE prote
11	702	36.0	413	2 F70560	probable PPE prote
12	688.5	35.3	463	2 C70931	probable PPE prote
13	682.5	35.0	468	2 G70932	probable PPE prote
14	675	34.6	380	2 A70846	probable PPE prote
15	668.5	34.3	394	2 G70881	probable PPE prote
16	667.5	34.2	385	2 H70503	probable PPE prote
17	635	32.6	350	2 H70929	probable PPE prote
18	633.5	32.5	365	2 H70929	probable PPE prote
19	615.5	31.6	402	2 A70882	probable PPE prote
20	603	30.9	423	2 C70882	probable PPE prote
21	597.5	30.7	391	2 D70922	probable PPE prote
22	589	30.2	406	2 G70675	probable PPE prote
23	588	30.2	394	2 A70504	probable PPE prote
24	575.5	29.5	391	2 A70663	probable PPE prote
25	495.5	25.4	3300	2 D70575	probable PPE prote
26	473	24.3	180	2 G70834	probable PPE prote
27	452.5	23.2	3716	2 G70969	probable PPE prote
28	448.5	23.0	580	2 G70570	probable PPE prote
29	444	22.8	346	2 H70874	probable PPE prote

30	443	22.7	963	2 B70524	probable PPE prote
31	441	22.6	487	2 C70830	probable PPE prote
32	439.5	22.6	655	2 A70931	probable PPE prote
33	429.5	22.0	479	2 D70676	probable PPE prote
34	428.5	22.0	1053	2 C70987	probable PPE prote
35	426.5	21.9	443	2 C70780	probable PPE prote
36	423.5	21.7	2523	2 F70846	probable PPE prote
37	422.5	21.7	615	2 E70663	probable PPE prote
38	418	21.4	678	2 A70762	probable PPE prote
39	418	21.4	3157	2 B70969	probable PPE prote
40	415.5	21.3	618	2 H70552	probable PPE prote
41	415	21.3	582	2 F70675	probable PPE prote
42	414.5	21.3	987	2 E70808	probable PPE prote
43	412	21.1	590	2 E70946	probable PPE prote
44	412	21.1	645	2 F70825	probable PPE prote
45	409.5	21.0	409	2 A70647	probable PPE prote

ALIGNMENTS

RESULT 1

B70608

Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: B70608

R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70608

A>Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-391 <CO>

A:Cross-references: GB:Z93777; GB:AL123456; NID:g3261726; PIDN:CAB07839.1; PID:e311073; I

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PPE

Query Match 100.0%; Score 1949; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 3.5e-108;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSARMYAGPGSASILVAAQWMDSVASDLFSAAGAFQSVWGLTVGSMIG 60

Db 1 MVDGALPEINSARMYAGPGSASILVAAQWMDSVASDLFSAAGAFQSVWGLTVGSMIG 60

QY 61 SSAGLWVAASPYVAVMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120

Db 61 SSAGLWVAASPYVAVMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120

QY 121 LIATNLLGNTPATVANEAEYGEWMAQDAAMFGVAATATATATLLPFEAPEMTSAGG 180

Db 121 LIATNLLGNTPATVANEAEYGEWMAQDAAMFGVAATATATATLLPFEAPEMTSAGG 180

QY 181 LLEQAAVAEEASDTAAANQLMNVPAQLQAQPTQGTTPSSKLGGLWKTYSRHSPI 240

Db 181 LLEQAAVAEEASDTAAANQLMNVPAQLQAQPTQGTTPSSKLGGLWKTYSRHSPI 240

QY 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQVCTAAQNGVRAMSLGSSLGSSG 300

Db 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQVCTAAQNGVRAMSLGSSLGSSG 300

QY 301 LGGGVAANLGRAASVGSLSVFPQAAANQAVTPAARALPLTSLTSAAPERGQMLGLPLV 360

Db 301 LGGGVAANLGRAASVGSLSVFPQAAANQAVTPAARALPLTSLTSAAPERGQMLGLPLV 360

QY 361 GOMGARACGGSLGVLRVPPRPYVMPHSPAAG 391

|||||

Qy	122	IATNLLGONTPAIAUNNABYGEWMAQDAAMFGYAAATATATATATLLFFPEAPEMWTSAGGL	181
Db	121	ISTNVFGNTSAIAAEAAQYGEWMAQDGAAMYAGSASASA-VTFSTPPOIANPTAQ	179
Qy	182	LEQAAAVEEASDTAAA--NQLMNNVFOALQQLAOPTQCTTTPSSKLGGLWK-----	229
Db	180	GTQAAAVATAAGTAQSTLTMTITGLPNALQSLTSPLOSS-NGPLSLWQLFSTPNFFT	238
Qy	230	-----TVSPHRSPISNNMVSVAANHMMWNTSGVSMNTLSSMLKGFAPAPAAAQAVQTAA	282
Db	239	SISALITDLPQYASFPYNTTEGLPYFSGVGNFNITQSAKTL-GLTGSAPAAVA-----AA	292
Qy	283	QNGVRAMSLGGSSLGSGGGVGAANLGRAASVGSLSVPOAWAA-ANQAVTPAARALPLT	341
Db	293	GDAAKGLPLGGMLG----GGPFAAGLGNAAASVGKLSVPPVWAGPLGSGVTPGNAALPVS	348
Qy	342	SLTSAAEERGPGQVLGGPLGVQWGAAGAGGSLGVLRVPRPPVMPHSPFAAG	391
Db	349	TVSAAPAEAPAGSILGGLPL---AGAGGAGAGP-RYGFRPTVMARPPFAG	393

RESULT 5

A70932

probable pPE protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: A70932

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Connor, R.; Davies, R.; Devlin, K.; Felkewell, T.; Gentles, S.; Hamlin, N.; H. H. Rajandream, M.A.; Rogers, R.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393, 537-544, 1998

A:Reference number: A70500; UID:98295987; PMID:9634230

A:Accession: A70932

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-409 <COL>

A:Cross-references: GB:AL022021; GB:AL123456; NID:G3250699; PIDN:CAA17729.1; P

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: pPE

Query Match	38.5%;	Score	751;	DB 2;	Length	409;			
Best Local Similarity	44.7%;	Pred.	No. 2.2e-37;						
Matches	177;	Conservative	53;	Mismatches	138;	Indels	28;	Gaps	10
<hr/>									
Qy	2	VDFGALPPEINSARMYAGPGSASLVAAQMDWSVASDLFSAAFSQFVWGLTVGSGWIGS	61						
Db	1	MDFGALPPEINSGRMTAGPGSGFLPLAAAAAADLAAEDLSAAASYGTITIGLTVAPWMPG	60						
Qy	62	SAGLMTVAASAPYVAMWSVTAGQAEILTAQVFAAAAAAYETVGLTVPFPVIAENRAELMIL	121						
Db	61	SSITMAAAVAVVAMISVTAGQEQAGQAKIAGVYETAFAATVPPVPIEANRALLMSL	120						
<hr/>									
Qy	122	IATNLIGQNTPAIVANEABYGEWQAODAAAMFGVAATAATATATATLLFFBEPAPENTSGAGL	181						
Db	121	VATNIFQNTPTAIAATEAHYAEMWQAODAAAMYGAGSSATA-SQLAPFSEPPOTTPPSAT	179						
<hr/>									
Qy	182	LEQAAAVEEASDTAAA-----NOLMKNVPQALQOLAQPTQGTTPSSKGLGLWKTVPSPH	234						
Db	180	AAGSANNVQAAGNAASDDITAQ:SQLISLLFSTLQSLA--TTATATGASAG--WDIV---232							
<hr/>									
Qy	235	RSPISNMVSNVNMHMSMTNSGVS-----MTNTLSMLKGFPAAAAAQVOTAAQNGVRAM	289						
Db	233	LQSITITILANTPGYSIIGLGAI PGCGWLLTFQQLGLAQNAPGVAALLGPKAAAGALSPL	292						
<hr/>									
Qy	290	SSL-GSSLGS-SGLGGVVAANI GRAASVGSISVPOQWNAANOQVTPAARALPITSLSA-	346						
Db	293	APURGGYIGDITPLGGGATGGTARAIYVGSISUSVPGWAEAAAPVNRVASVLPGTGAAPAL	352						
<hr/>									
Qy	347	AERPGPQMLGGUPVGMQGARAGCGGL-----SGVLRV	377						

Df 353 AAEAPGALFGEVALSSLAGRALAGTAVRSGAARV 388

RESULT 6
H70931
Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: H70931
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gencies, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Scates, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: H70931
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-403 <COL>
A: Cross-references: GB: AL020201; GB: AL123456; NID: G3250699; PIDN: CAAL1728.1; PID: e125461
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: PPE

Query Match 37.8%; Score 737.5; DB 2; Length 403;
Best Local Similarity 42.6%; Pred. No. 1.4e-36;
Matches 181; Conservative 60; Mismatches 123; Indels 61; Gaps 12;

Qy 2 VDFGLPPEINSARMYAGPGSASLVAAQMDVSADLPSAASAFOSVVMGLTVGVWIGS 61
Db :
5 LDFATLPPEINSARMYAGGAPMLAASAMHGLSAELRSALSYSVLSTLTGEEWHGP 64
Qy 62 SAGLMVAASPYPVAMWSVTAGQAEHTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMIL 121
Db :
65 ASASNTAAAPVYVAMWSVTAVREAQAQAQAEAAAAAYEAFAATVPPIEANRAQLMAL 124
Qy 122 IATNLGONTAIIVNEAEYGEHMWAODAAAMEGYAANTATATATLTPFEAPEWTSAGGL 181
Db :
125 IATNVLGQNAPAIATEAQYAEMWSODAMTYGAGASAAAT-QLTPTFTPEVOTTNASGL 183
Qy 182 LEQAAVBEASDTEAAA-----NQLMNNVPOALQOLAQPQTGT-----TPSS 222
Db :
184 AAQSAIIAHATGASAGAQQTTLSQLIAAIPSVLQGLSSSTAATFAFGSPGSLGIVGSGSS 243
Qy 223 KLGLWKTVSPHRSPIINMVSNANDHEMSWTNSGVSM-TNTLS---SWLKGFAPAARAAQAV 278
Db :
244 WLDKXLWALLDEN-----SNFWNTIASGGLFLPNSNTIAPFLGLLGGVAAADAAGDV 293
Qy 279 QTAQNGVRAMSSSLGSSL-----GSSGLGGGVAAAMLGRAASVGSLSVPPOAWAANAQAVTP 333
Db :
294 LGKATSG----GLGGVALVPLGSGGLGGTVAAGLGNAAATVGTLSPVPSWTAAPLASP 348
Qy 334 AARAL---PLTSLTISAERGPQMGLGGLPVGMQBARAGGLSGVLRVPP-----RPYVMPH 386
Db :
349 LGSALGGTFPMVAPPVPAVAG-----MPCMFPTMGQGQFG-----RAVPOYGRFPNPVAR 398
Qy 387 SPAAG 391
Db : : : : :
399 PPAG 403

RESULT 7
B70931
Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: B70931
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gencies, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

```

RESULT 8
H87056
PPF-family protein [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87056
R:Colet, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hsiao, C.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, J.; Kaur, M.A.; Rutherford, K.M.
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86309; MUID:21128732; PMID:11234002
A:Accession: H87056
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-421 <STO>
A:Cross-references: GB:AL450380; NID:g13093150; PIDN:CAC31563.1; GSPDB:GNC0147
C:Genetics:
A:Gene: M1182

Query Match 37.5%; Score 731; DB 2; Length 421;
Best Local Similarity 41.1%; Pred.No. 3.5e-36;
Matches 174; Conservative 55; Mismatches 160; Indels 34; Gaps 8;

```

```
Db 180 AGQAATVSTVPLPATTAAVPLQQLSLSLIPWYSLQQLWLAENLLGLTPDNMTIV 239
QY 225 -----GGLWKTSPHRSPTSNMVMANNHMTNSGVSMTNTLSMLKGFAPAAAA 275
Db 240 RLLGISYFDEGL-----LQFEASLAQAIPGTPGAG--DSGSSVLDSWGPTIFA 287
QY 276 QAVCTAAQNGVRWSSL--GSSLSG-----SGLGGVAAVLGRAASVGSLS 319
Db 288 -----GPRASPVSAGGAVGGVQTPOPYWYWDRESIGSVAALGKSSAGSLS 338
QY 320 VPOWAAANQAVTPAARALP--LTSLSAERPGCMGLGLPVQGMGARAGGSLVLR 376
Db 339 VPDWAAARAWAPAWPLPDGDVTLALRGTAENA-----LURGFMASAGSTGGGF--VHK 393
QY 377 VPPRPYMPHSPAAG 391
Db 394 YGFLAWMPRPFPAG 408

RESULT 10
B70625
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70625
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70625
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-391 <COL>
A:Cross-references: GB:Z92539; GB:AL123456; NID:G3261714; PIDN:CAB0687.1; PID:e304546;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PPE
```

```
Query Match 36.0%; Score 702; DB 2; Length 391;
Best Local Similarity 42.3%; Pred. No. 1.6e-34;
Matches 166; Conservative 62; Mismatches 150; Indels 14; Gaps 7;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIGS 61
Db 1 MDFGALPPEINSARMYAGAGPMWAGAAWGLAAELGTTAASVETRLTTESWMGP 60
QY 62 SAGLMVAASPYVAMSVTAGQAELETAQVRAAAAYETAYGLTVPPVPIAENRAELMLI 121
Db 61 ASMAVAAAPYLAWLTYTEAAHAGSQAASAAAYEAYAMTVPEVVAANRALLAAL 120
QY 122 IATNLLGQNTPAIAVNEAEGENWAOQAAAFYAAATATATATLLPFEAPEMTSAGGL 181
Db 121 VATNVLGINTPAIMATEALYAENWAOQALAMYGAAASG-AAQMLQPLSPSPQTTNPGGL 179
QY 182 LEQAAAVEASDSTAANQ-----LMNVPOALQQLAQTOGTTPTPSKLGGLWKTSPHRS 236
Db 180 AAGSAVGSAAATAVNVQSVADLSLPAVNSGLASPTSVLSDTGLSGIADIDALLA 239
QY 237 P--ISNMVSNVNNHMS--MTNSGVSMTNTLSMLKGFAPAAAAQAVTAQNGVRWAMSSIG 293
Db 240 TPFVANIINSVNTAAWYVNAAIPTAIFLANALNSGAPVAIEGAIEAAEG--AASAAA 296
QY 294 SFLGSSGLGGVAAVTGRAASVGSLSVPOWAAANQAVTPAARALPLTSLTSAERGPQ 353
Db 297 AGLADSVTPAGLGSAGEATLVRSLVSPAAMWSTAAAPTATAGATALEGSWTVAAEEA-CP 355
QY 354 MLGGLPVQGMGARAGGSLGVLRVPPRPYMP 385
Db 356 VTGMWP-GWASAAKGTGAVAGPRYGFKPTWMP 386.
```

RESULT 11

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: F70560

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70560

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-413 <COL>

A:Cross-references: GB:Z95436; GB:AL123456; NID:G3261770; PIDN:CAB08826.1; PID:e316565; I

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: PPE

Query Match 36.0%; Score 702; DB 2; Length 413;

Best Local Similarity 39.1%; Pred. No. 1.7e-34;

Matches 168; Conservative 61; Mismatches 145; Indels 56; Gaps 7;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60

Db 1 MLDFPALPPEVNSALMYAGPGSGPMLAAAAAAWEALAELOTTASTYDALITGLADGPWQG 60

QY 61 SSAGLMVAASPYVAMSVTAGQAELETAQVRAAAAYETAYGLTVPPVPIAENRAELMI 120

Db 61 SSAASVMAAATPQVAAWLRSTAGQAEQAGSQVAAASAYEAFAFFATVPPEIAANRALLMA 120

QY 121 LIATNLLGQNTPAIAVNEAEGENWAOQAAAFYAAATATATATLLPFEAPEMTSAGG 180

Db 121 LLATNLFQNTAAIAATEAQVRAAAAYETAYGLTVPPVPIAENRAELMI 179

QY 181 LLEQAAAVEASDSTAANQLMNVPOALQQLAQTOGTTPTPSKLGGLWKTSPHRSISP 240

Db 180 LASQAASVQAVSGVGAANAQAALTDIPKAL-----FGLSGIFNPPMLTDLGLK 226

QY 241 MVSMAANNHMTNSGVSMTNTLSMLKGFAPAA--AAQAVOT----- 280

Db 227 ALGLTCHTWSSDGGSLIVGGVIGDFVQCVTSABELDASVMDTFKXVSPARLMTQPKD 286

QY 281 -----AAQNGVRAMSSLSGSSGLGGVGAANGLGRAASVGSLSVPOA 323

Db 287 YFGLAHDLPKWASBGAAGAARAKALPAAPVPAIPSAGL--SGVAGAVGQAASVGGGLKVPV 345

QY 324 WAAANQAVTPAARALPLTSLTSAERGPQMLGGLPVQGMGARAGGSLGVL--RVPRP 381

Db 346 WTATTPAASPVAALASNLGAAABEGSTHAFGGMLP--MGSGAGRAFNFAPRYGFKP 403

QY 382 YMPHSPAAG 391

Db 404 TVIAQPPAGG 413

RESULT 12

probable PPE protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: C70931

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

Db 290 GAASVGNVTILASVGRANSIGQLSVPPSWAAPSTRPVLSALSPAGLTLTLPCTDVAEHGMPG- 348
QY 354 MLGGLPVGQMGARAGGLSGVL-RVPPRPVYVMPHSPAAG 391
Db 349 -VPGVPV-----AAGRASGVLPYGVRLTVMAHPPAAG 380

RESULT 15
G70881
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
R;Accession: G70881
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Squires, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: G70881
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-394 <COL>
A;Cross-references: GB:AL008967; GB:AL123456; NID:g3261491; PIDN:CAA15564.1; PID:e117389
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: PPE

Query Match 34.3%; Score 668.5; DB 2; Length 394;
Best Local Similarity 40.8%; Pred. No. 1.6e-32;
Matches 161; Conservative 61; Mismatches 156; Indels 17; Gaps 7;

QY 2 VDFGALPPEINSRMYAGSGSLVAAAQWDSVASDLFSAASAFOSVYVWGLTVGSGWGS 61
Db 1 MDFGALPPEINSTRMYAGAGAPLMAAGATNGLAVELSTASSVESVIMQLTTEQWLGP 60
QY 62 SAGLWYAAASPYVAMSVTVAGQELTAAQVRVAAAVETAYGLTVPPPVIAENRAELMIL 121
Db 61 ASMWVYAAQPYLAWLTYTHESAHAHAQAQMAAFAFAFAMTVPAEVANRALLAAL 120
QY 122 IATNLLGQNTPAIVNNEAEYGEWMAQDAAMFCYAAATATATATLFPFEAPEMTSAGGL 181
Db 121 VATNVLGQNTPAIWAIEAHYGEWMAQDALAMYAAASSA-AAGRLNPLITPSTQANMAGL 179
QY 182 LEQAAAVEEASDTAAHQ-----LMNVFQALQQLAQPTGTTTPSSKIGLWKTVP--H 234
Db 180 AGQAAAVSHAAASTVQQVGLGLISNLPNAVMGFASPLTSAADAAGLGGITQDIEELIG 239
QY 235 RSPISNVSMANNHMSMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQNGV---RAMSS 291
Db 240 ITFQNAINGAVN-----TTAFVYMATIPNAVFLGHAFALNPATVTAADAVPAAAAAAG 295
QY 292 LGSLSGSSGLGG-GVAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERG 350
Db 296 LAHTVTPVGVGASLTASLGEASSVGLSVPGAGWSTAAPAMTSGTTALEGSGWAVPEEAG 355
QY 351 PGQWMLGGLPVGQMGARAGGLSGVLVPPRPVYMP 385
Db 356 PVAAMPGN-AGISGAAGKAGAYAGPYGPKPIVNP 389

Search completed: August 25, 2004, 03:13:46
Job time : 41 secs

This Page Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:05:03 ; Search time 118 Seconds

(without alignments)

1045.488 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDFGALPEINSARMYAGP.....SGVLVPPRPVYMPHSPAAG 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1949	100.0	391	16	C05298 mycobacteri
2	1920.5	98.5	390	16	Q7U0E9
3	1779.5	91.3	396	16	Q7U071
4	1648	84.6	393	16	Q8VIZ3 mycobacteri
5	1583	81.2	393	16	C06341 mycobacteri
6	1583	81.2	393	16	Q7TW65 mycobacteri
7	779.5	40.0	393	16	Q7TZJ3 mycobacteri
8	779.5	40.0	411	16	O53939 mycobacteri
9	754	38.7	410	2	Q99QI1 mycobacteri
10	751	38.5	409	16	O53957 mycobacteri
11	749	38.4	409	16	Q7TZH7 mycobacteri
12	740.5	38.0	399	16	Q7TZH8 mycobacteri
13	737.5	37.8	403	16	O53956 mycobacteri
14	736	37.8	423	16	O53950 mycobacteri
15	736	37.8	423	16	Q7TZI4 mycobacteri
16	731	37.5	421	16	Q9Z5K0 mycobacteri

17	702	36.0	391	16	P96352	P96352 mycobacteri
18	702	36.0	391	16	Q7U0E5	Q7U0E5 mycobacteri
19	702	36.0	413	16	O06386	O06386 mycobacteri
20	683.5	35.1	694	16	Q8VJW0	Q8VJW0 mycobacteri
21	682.5	35.0	468	16	O53958	O53958 mycobacteri
22	675	34.6	380	16	P95190	P95190 mycobacteri
23	675	34.6	380	16	Q7TX66	Q7TX66 mycobacteri
24	670.5	34.4	385	16	Q7TZ87	Q7TZ87 mycobacteri
25	669.5	34.4	394	16	Q7TX5	Q7TX5 mycobacteri
26	668.5	34.3	462	16	O33310	O33310 mycobacteri
27	667.5	34.2	385	16	O33204	O33204 mycobacteri
28	666.5	34.2	385	16	Q8VJZ0	Q8VJZ0 mycobacteri
29	641	32.9	364	16	Q7TZ35	Q7TZ35 mycobacteri
30	637	32.7	350	16	Q7TZ32	Q7TZ32 mycobacteri
31	635	32.6	363	16	O53940	O53940 mycobacteri
32	633.5	32.5	365	16	O86373	O86373 mycobacteri
33	629.5	32.3	405	16	Q8VJW5	Q8VJW5 mycobacteri
34	626.5	32.1	381	16	Q7TX67	Q7TX67 mycobacteri
35	618.5	31.7	397	2	Q9AGF0	Q9AGF0 mycobacteri
36	615.5	31.6	382	16	Q7TXX3	Q7TXX3 mycobacteri
37	615.5	31.6	402	16	O33312	O33312 mycobacteri
38	610	31.3	443	16	Q8VKL9	Q8VKL9 mycobacteri
39	604	31.0	443	16	Q7U242	Q7U242 mycobacteri
40	603	30.9	423	16	Q7U114	Q7U114 mycobacteri
41	603	30.9	426	16	O05907	O05907 mycobacteri
42	597.5	30.7	391	16	O05798	O05798 mycobacteri
43	594.5	30.5	391	16	Q7TX76	Q7TX76 mycobacteri
44	589	30.2	406	16	P71869	P71869 mycobacteri
45	589	30.2	406	16	Q7TW99	Q7TW99 mycobacteri

ALIGNMENTS

RESULT 1: 100%
005298 PRELIMINARY; PRT; 391 AA.
ID C05298
AC C05298;
DT 01-JUL-1997 (TREMREL. 04, Created)
DT 01-JUL-1997 (TREMREL. 04, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Hypothetical protein (PPE family protein).
GN RV1196 OR MTC1364.08 OR MT1234.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.B. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutten S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayama L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bisshai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.


```

QY 181 LLEQAAVEASDTAAANLNNVPOALQOLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
Db 181 LLEQAAVEASDTAAANLNNVPOALQOLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
QY 241 MVMANNHMTNSGVSMTNTLSMLKGFAPAAAAQAVTAQAQNGVRMSS-----LGSSL 296
Db 241 IVSMLNHNHVSMTNSGVSMTNTLSMLKGFAP-AAAQAVETAQNGVQVQVMSLSLGS 299
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRPQMLG 356
Db 300 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRPQMLG 359
QY 357 GLPVGQMGARAG--GGLSGVLRVPPRPVPMHPSPAAAG 391
Db 360 GLPLGLQNTSGCGFGGVSNALRMPRPAYVMPVPAAG 396

RESULT 4
Q8VIZ3
ID Q8VIZ3 PRELIMINARY; PRT; 393 AA.
AC Q8VIZ3;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE PPE family protein.
GN MT3582.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Dalcier A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007161; AAK47941.1; -.
DR TIGR; MT3582; -.
DR GO; GO:0016020; C.membrane; IEA.
DR GO; GO:0008233; F.peptidase activity; IEA.
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro; IPR000030; Microbac_PPE.
DR InterPro; IPR000508; Peptidase_S26.
DR Pfam; PF00823; PPE; 1.
DR PROSITE; PS00501; SPASE_I_1; 1.
SQ SEQUENCE 393 AA; 39658 MW; 86F0B67798855511 CRC64;

Query Match 84.6%; Score 1648; DB 16; Length 393;
Best Local Similarity 84.8%; Pred. No. 2e-85;
Matches 335; Conservative 20; Mismatches 34; Indels 6; Gaps 3;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWMDVSADLPSAASAFQSVVGLTVGSGWIG 60
Db 1 MVDFGALPPEINSARMYAGPGSASLVAAQWMDVSADLPSAASAFQSVVGLTVGSGWIG 60
QY 61 SSAGLMVAASPYVAMWSVTAGAEELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
Db 61 SSAGLMVAASPYVAMWSVTAGAEELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
Db 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
QY 181 LLEQAAVEASDTAAANLNNVPOALQOLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
Db 181 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
QY 241 MVMANNHMTNSGVSMTNTLSMLKGFAPAAAAQAVTAQAQNGVRMSS-----LGSSL 296
Db 241 LLEQAAVEASDTAAANLNNVPOALQOLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRPQMLG 356

```

```

Db 241 IVSMLNHNHVSMTNSGVSMTNTLSMLKGFAP-AAAQAVETAQNGVQVQVMSLSLGS 299
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRPQMLG 356
Db 300 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRPQMLG 359
QY 357 GLPVGQMGARAGGGLSGVLRVPPRPVPMHPSPAAAG 391
Db 360 GLPLGH-SVNAAGSINNLRVPAAYALPRTPAAG 393

RESULT 5
Q06341
ID Q06341 PRELIMINARY; PRT; 393 AA.
AC Q06341;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein RV3478.
GN RV3478 OR MTCY13E12.31.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Healey S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
DR EMBL; Z95390; CAB08702.1; -.
DR PIR; C70568; C70568.
DR Tuberculist; RV3478; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 393 AA; 39413 MW; AF4C20C95DAE7DD4 CRC64;

Query Match 81.2%; Score 1583; DB 16; Length 393;
Best Local Similarity 81.5%; Pred. No. 9.3e-82;
Matches 322; Conservative 20; Mismatches 47; Indels 6; Gaps 3;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWMDVSADLPSAASAFQSVVGLTVGSGWIG 60
Db 1 MVDFGALPPEINSARMYAGPGSASLVAAQWMDVSADLPSAASAFQSVVGLTVGSGWIG 60
QY 61 SSAGLMVAASPYVAMWSVTAGAEELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
Db 61 SSAGLMVAASPYVAMWSVTAGAEELTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEAPEMTSAGG 180
Db 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEAPEMTSAGG 180
QY 181 LLEQAAVEASDTAAANLNNVPOALQOLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
Db 181 LLEQAAVEASDTAAANLNNVPOALQOLAQPTGTTTSSKLGGLWKTVPSPHRSPI 240
QY 241 MVMANNHMTNSGVSMTNTLSMLKGFAPAAAAQAVTAQAQNGVRMSS-----LGSSL 296
Db 241 VSSIANNHMTNSGVSMTNTLSMLKGLAP-AAAQAVETAQNGVQVQVMSLSLGS 299
QY 297 GSSGLGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTAAAGRPQMLG 356

```

```

Db 300 GSSGLGAGVAANLGRAASVGSLSVPPFAWAANQVTPAARALPLTSLTSAQAQAPGHMLG 359
QY 357 GLPVQGVQAGAGGSLGVLVPPRPVPMHPSAAG 391
Db 360 GLPLGH-SVNAGSGINNLRVAPARAIVAPRTPAAG 393

RESULT 6
Q7TWFS PRELIMINARY; PRT; 393 AA.
AC Q7TWFS, MEDLINE=22709107; PubMed=12788972;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PE family protein.
GN PPE60 OR MB3505.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteriae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=AF2122/97;
RC MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248346; CAD95692.1; -.
KW Complete proteome.
SQ SEQUENCE 393 AA; 39413 MW; AF4C20C95DAE7DD4 CRC64;

Query Match 81.2%; Score 1583; DB 16; Length 393;
Best Local Similarity 81.5%; Pred. No. 9.3e-82;
Matches 322; Conservative 20; Mismatches 47; Indels 6; Gaps 3;

QY 1 MVDFGALPPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60
Db 1 MVDFGALPPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60
QY 61 SSAGLWVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
Db 61 SSAGLWVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLTPPEAPEMTSAGG 180
Db 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLTPPEAPEMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNVPOALQOAOPTGTTTSSKLGGLWKTVPSPHSPISN 240
Db 181 LLEQAAVEEASDTAAANQLMNVPOALQOAOPTGTTTSSKLGGLWKTVPSPHSPISN 240
QY 241 MVSMANNHMTNSGVSMNTLLSSMLKGFAPAQAQAVCTAAGQVRAMSS---LGSSL 296
Db 241 VSSIANHMSMTGTVSMNTLLSHMLKGLAP-AAQAVETAENGVAMSSLSGSSL 299
QY 297 GSSGLGGVGAANLGRAASVGSLSVPPQAAANQAVTPAARALPLTSLTSAARPGQMLG 356
Db 300 GSSGLGAGVAANLGRAASVGSLSVPPFAWAANQVTPAARALPLTSLTSAQAQAPGHMLG 359
QY 357 GLPVQGVQAGAGGSLGVLVPPRPVPMHPSAAG 391
Db 360 GLPLGH-SVNAGSGINNLRVAPARAIVAPRTPAAG 393

RESULT 7
Q7TZJ3 PRELIMINARY; PRT; 393 AA.
ID Q7TZJ3
AC Q7TZJ3, MEDLINE=98295987; PubMed=9634230;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PPE family protein.
GN PPE26 OR MB1817.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteriae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=AF2122/97;
RC MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248340; CAD94520.1; -.
KW Complete proteome.
SQ SEQUENCE 393 AA; 38572 MW; 41944B6B547A8A80 CRC64;

Query Match 40.0%; Score 779.5; DB 16; Length 393;
Best Local Similarity 43.7%; Pred. No. 2e-36;
Matches 179; Conservative 65; Mismatches 129; Indels 37; Gaps 10;

QY 2 VDFGALPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIGS 61
Db 1 MDFGALPEVNSVEMYAGPGSAPVVAASAWNGLAELSSAATGYETVITQLSSEGWLGP 60
QY 62 SAGLWVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 121
Db 61 ASAMAASAVAPYVAMSGAAQAQAAQAAQAAQAAQAAQAAQAAQAAQAAQAAQAA 120
QY 122 IATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLTPPEAPEMTSAGG 181
Db 121 ISTNVFGQNTSAIAAAEAQYGEWMAQDAAMFGYAAATATATATLTPPEAPEMTSAGG 179
QY 182 LEQAAVEEASDTAAANQLMNVPOALQOAOPTGTTTSSKLGGLWKTVPSPHSPISN 229
Db 180 GTQAAVATAAGTAQSTLTITMITGLPNALQSLTSPLLQSS-NGPLSLWQLFQTPNFPT 238
QY 230 -----TVSPHSPISNMVSMANNHMTNSGVSMNTLLSSMLKGFAPAQAQAVCTAA 282
Db 230 -----TVSPHSPISNMVSMANNHMTNSGVSMNTLLSSMLKGFAPAQAQAVCTAA 282
QY 239 SISALLTDQYAGFFNYTEGLPYFSIGMGNFIQAAKTL-GLIGSAAPAAVA-----AA 292
Db 239 SISALLTDQYAGFFNYTEGLPYFSIGMGNFIQAAKTL-GLIGSAAPAAVA-----AA 292
QY 283 QNGVRAMSSLSGSSLSGSSGLGGVGAANLGRAASVGSLSVPPQAAANQAVTPAARALPLT 341
Db 293 GDAAKGLPGLGMLG---GGPVAAGLGNAAASVGKLSVPPVMSGSLPGSVTPGAAPLVS 348
QY 342 SLTSAARPGQMLGGLPVGMQGARAGGSLGVLVPPRPVPMHPSAAG 391
Db 349 TVSAAPAAPGSLGGLPFL-----AGAGGAGGP-RYGFRTVMARPPFAG 393

RESULT 8
OS3939 PRELIMINARY; PRT; 411 AA.
ID OS3939
AC OS3939, MEDLINE=98295987; PubMed=9634230;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PPE-family protein.
GN RV1789 OR MT1838 OR MTV049.11.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteriae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=H37Rv.
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

```

RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022021; CAA17711.1; ALT_INIT.
DR EMBL; AB007043; AAK46108.1; -;
DR PIR; G70929; G70929.
DR TIGR; M1838; -;
DR TubercuList; Rv1789;
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 411 AA; 40557 MW; 03C90B5E0590B7DA CRC64;

Query Match 40.0%; Score 779.5; DB 16; Length 411;
Best Local Similarity 43.7%; Pred. No. 2.1e-36;
Matches 179; Conservative 65; Mismatches 129; Indels 37; Gaps 10;

QY 2 VDFGALPPEINSARMYAGPGSALVAAAQWDSVADLSFAASAFQSVVMGLTVGSWIGS 61
DB 19 MDFGALPPEVNSVRMYAGPGSAPWVAAGSAWNGLAELSSAATGYETVITQLSSEGLGP 78
QY 62 SAGLWVAASPYVAVNSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVPIAENRAELMI 121
DB 79 ASAAAEAVAPYVAVNSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVPIAENRAELMI 138
QY 122 IATNLGQNTPAIVNEAEYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGL 181
DB 139 ISTNVFGQNTSAIAAEAYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGL 197
QY 182 LEQAAVEEASDTAAA--NOLMNNVPAQLQQLAQPTQGTTPSSKLGGLWK----- 229
DB 198 GTQAAAVATAAGTAQSTLTEMITGLFNALQSLTSPLLQSS-NGPLSLWQLLFTGFNFFT 256
QY 230 -----TVSPHSPISNNVSMANNHMTNSGVSMTNTLSSMLKGFAPAAAQAQVATAA 282
DB 257 SISALLTDQPIYASFFYNTGELPYFSGMGNFIQSAKIL-GLIGSAAPAAVA----- 310
QY 283 QNGVRAMSSISGLSSGLGGVAAANLGRAASVGSLSVFPQAWAA-ANQAVTPAARALPLT 341
DB 311 GDAAKGLPGLGGMLG---GGPVAAGLGNAAASVGLSVPPVWMSGPLPGSVTPGAAPLPVS 366
QY 342 SLTSAERGPQMLGLPVGMQCARAGGSLGSLVLRVPPPRPYMPHSPAAG 391
DB 367 TVSAPEAPAPGSLGLGLPL----AGAGGAGAGP-RYGFRPTVNRPPFAG 411

RESULT 9
Q99Q11 PRELIMINARY; PRT; 410 AA.
AC Q99Q11;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Rv1808-like protein.
GN MYC1808 OR OV1808.

OS Mycobacterium microti.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1806;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MYC 94-2272, and OV254;
RA Stravon V., Heym B., Mazancourt P., Gaillard J.-L.L.;
RT "PPE Rv1808 orthologue of Mycobacterium microti.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335180; AAK20894.1; -;
DR EMBL; AF335179; AAK20893.1; -;
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
SQ SEQUENCE 410 AA; 40044 MW; 3B9DA3174655A5EA CRC64;

Query Match 38.7%; Score 754; DB 2; Length 410;
Best Local Similarity 44.5%; Pred. No. 5.7e-35;
Matches 177; Conservative 51; Mismatches 140; Indels 30; Gaps 10;

QY 1 MVDGALPPEINSARMYAGPGSALVAAAQWDSVADLSFAASAFQSVVMGLTVGSWIG 60
DB 1 MLDFGALPPEINSARMYAGPGSGLLAAAADWDAALAAELYSAAASYGSTIEGLTVAPMNG 60
QY 61 SSAGLWVAASPYVAVNSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVPIAENRAELMI 120
DB 61 PSSITMAAAVAPYVAVNSVTAGQAEHTAAQVRVAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLGQNTPAIVNEAEYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGL 180
DB 121 LVATNIFGQNTPAIVNEAEYGEWMAQDAAMFYAAATATATATLPPFEAPEMTSAGL 179
QY 181 LLEQAAVEEASDTAAA-----NOLMNNVPAQLQQLAQPTQGTTPSSKLGGLWKTVSP 233
DB 180 TAQSAVVAQAAGAAASDITQAQLSLILPSTLQSLA--TTATATASAG--WDTV-- 233
QY 234 HRSFISNNVSMANNHMTNSGVS---SMTNTLSSMLKGFAPAAAQAQVATAAQNQVRAMS 290
DB 234 -LQSITITLANTLGPYSITGLGAIPTGGWLLTFGQIL-GLAQNAPGVAAALLGPXAAAGALS 291
QY 291 SIGSSLGS-----SGLGGVAAANLGRAASVGSLSVFPQAWAAANQAVTPAARALPLTSLTS 345
DB 292 FLAPLRGGYIADITPLGCGGATGGIARATYVGLSVPGQWAEAPVPMRAVASVLPFGTGAAP 351
QY 346 A-AERGPQMLGLPVGMQCARAGGGL-----SGVLRV 377
DB 352 ALAAEPALFGENALSSLAGRALAGTAVRSAGAARV 389

RESULT 10
O53957 PRELIMINARY; PRT; 409 AA.
AC O53957;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE PPE-family protein.
GN Rv1808 OR MT1856.1 OR MTV049.30.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RL complete genome sequence.";
 RN Nature 393:537-544(1998).
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RL laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL020221; CAA17729.1; -;
 DR FIR; AEO07044; AAK46129.1; ALT_INIT.
 DR TIGR; MT1856.1; -;
 DR TubercuList; Rv1808; -;
 DR InterPro; IPR000030; Microbac_PPB.
 DR Pfam; PF00823; PPE; 1.
 KW Complete proteome.
 SQ SEQUENCE 409 AA; 39917 MW; 1E15202BACF36379 CRC64;
 Query Match 38.5%; Score 751; DB 16; Length 409;
 Best Local Similarity 44.7%; Pred. No. 8.4e-35;
 Matches 177; Conservative 53; Mismatches 138; Indels 28; Gaps 10;
 QY 2 VDFGALPPEINARMYAGPGSASLVAAAQWDSVADLFSASAFQSVVWGLTVGWSIGS 61
 DB 1 MDFGALPPEINSGRMVAGPGSGPLLAALAAWDAALAEALYSAAASVGTIEGLTVAPWMPG 60
 QY 62 SAGLMVAASAPYVAMVSVTAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMIL 121
 DB 61 SSITMAAAYVAVVAVISVTAGQEQAGAKIAGVYETAFATVPPVIAENRALLMSL 120
 QY 122 IATNLGQNTPAIAVNEAEYGENWQAQDAAMFGYAAATATATATLFFPEAPMTSAGGL 181
 DB 121 VATNIFQNTPAIAATEAHYAEMWQAQDAAMYGAGSSATA-SQLAFFSPPTNPST 179
 QY 182 LEQAAVEEASDTAAA-----NQLMNVPOALQOLAQTOGTTPSSKLGGLWKTUSPH 234
 DB 180 AAQSAVVAAQAGAAASDITAQLSQLISLPLSTLQSLA--TTATATSASAG--WDTV--- 232
 QY 235 RSPISNMVSMANNHMTNSGV-----MTNTLSMLKGFAPAAAQAQVQTAQNGVRAM 289
 DB 233 LQSIITILANTGPYSIIIGLAIPEGWMLTFGQILGLAQNAFCVALLGPKAAGALSPL 292
 QY 290 SSL-GSSLGS-SGLGGVAAANLGRAASVGSLSVPOQAWAANQAVTPAARALPLTSLSA- 346
 DB 293 APLRGVIGDITPLGGGATGIARAIYVGSLSVPOGWAEEAAPVMRAVASVLEFGTGAAPAL 352
 QY 347 AERGGOMLGGPLVPGWGARGAGGL-----SGVLRV 377
 DB 353 AAAPGALFEMALSSLAGRALAGTAVRSGAGARV 388
 RESULT 11
 QY7TZH7
 ID QY7TZH7 PRELIMINARY; PRT; 409 AA.
 AC QY7TZH7;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE PPE family protein.
 GN PPE32 OR MB1837.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248340; CAD94540.1; -;
 KW Complete proteome.
 SQ SEQUENCE 409 AA; 39931 MW; D57892628B131A9E CRC64;
 Query Match 38.4%; Score 749; DB 16; Length 409;
 Best Local Similarity 44.3%; Pred. No. 1.1e-34;
 Matches 176; Conservative 51; Mismatches 140; Indels 30; Gaps 10;
 QY 2 VDFGALPPEINARMYAGPGSASLVAAAQWDSVADLFSASAFQSVVWGLTVGWSIGS 61
 DB 1 MDFGALPPEINSGRMVAGPGSGPLLAALAAWDAALAEALYSAAASVGTIEGLTVAPWMPG 60
 QY 62 SAGLMVAASAPYVAMVSVTAGQELTAQVRVAAAYETAYGLTVPPVIAENRAELMIL 121
 DB 61 SSITMAAAYVAVVAVISVTAGQEQAGAKIAGVYETAFATVPPVIAENRALLMSL 120
 QY 122 IATNLGQNTPAIAVNEAEYGENWQAQDAAMFGYAAATATATLFFPEAPMTSAGGL 181
 DB 121 VATNIFQNTPAIAATEAHYAEMWQAQDAAMYGAGSSATA-SQLAFFSPPTNPST 179
 QY 182 LEQAAVEEASDTAAA-----NQLMNVPOALQOLAQTOGTTPSSKLGGLWKTUSPH 234
 DB 180 AAQSAVVAAQAGAAASDITAQLSQLISLPLSTLQSLA--TTATATSASAG--WDTV--- 232
 QY 235 RSPISNMVSMANNHMTNSGV-----MTNTLSMLKGFAPAAAQAQVQTAQNGVRAMSS 291
 DB 233 LQSIITILANTGPYSIIIGLAIPEGWMLTFGQILGLAQNAFCVALLGPKAAGALSPL 291
 QY 292 LGSLSG-----SGLGGVAAANLGRAASVGSLSVPOQAWAANQAVTPAARALPLTSLSA 346
 DB 292 LAPLRGGVIADITPLGGGATGIARAIYVGSLSVPOGWAEEAAPVMRAVASVLEFGTGAAPA 351
 QY 347 -AERGGOMLGGPLVPGWGARGAGGL-----SGVLRV 377
 DB 352 LAAPGALFEMALSSLAGRALAGTAVRSGAGARV 388
 RESULT 12
 QY7TZH8
 ID QY7TZH8 PRELIMINARY; PRT; 399 AA.
 AC QY7TZH8;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE PPE family protein.
 GN PPE31 OR MB1836.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248340; CAD94539.1; -;
 KW Complete proteome.
 SQ SEQUENCE 399 AA; 38840 MW; 1A0F4377318E74F2 CRC64;
 Query Match 38.0%; Score 740.5; DB 16; Length 399;


```

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Unwayan L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL020201; CAAl7722.1; -.
DR EMBL; AE007044; AAK46122.1; -.
DR PIR; B70931; B70931.
DR TIGR; MT1850; -.
DR TubercuList; Rv1801; -.
DR InterPro; IPR00030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
DR Complete proteome.
SQ SEQUENCE 423 AA; 41477 MW; 26E52CC271FBBF57 CRC64;

Query Match 37.8%; Score 736; DB 16; Length 423;
Best Local Similarity 41.8%; Pred. No. 6.2e-34;
Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;

QY 2 VDFGALPPEINARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGWSIGS 61
DB 1 MDFGLLPPEINSGRMVTPGPGPMLAAATAWDGLAVELHATAAGYASELSALT-GAWSGP 59

QY 62 SAGLMVAASPYVAVMSVTAGQAELETAQVRAAAVETAYGLTVPPVIAENRAELMIL 121
DB 60 SSTMSASAAAPYVAVMSATVHAELAGAQARLAIAAYEAFAATVPPVIAANRAQLMVL 119

QY 122 IATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGGL 181
DB 120 IATNIFGQNTPAIWMTEAQYEMWMAQDAAMFYAGSSATA-SRMTAFTEPPQTTHGQL 178

QY 182 LEQAAVEEASDTAAAN-----QLMNVVPAQLQLOAQPT-----QGTFP-----S 221
DB 179 GAQSSAVAQTAATAAGNLOSAPFQLLSAVPRALQGLALPTASQSASATPQWVTDLGNLS 238

QY 222 SKLGLMKTVPSPHRSPTSNMVMNMHMTNSGVSMTLSSMLKGFAPAAAAAQVTA 281
DB 239 TFLGG--AVTGPYTFP-----GVLPPSGVPYLLGIQSVL-----V 271

QY 282 AQNGVRAMSSILGS-----SLGSSGLGGG--VAANLGRAASVGSLS 319
DB 272 TQNGQGVSAALLGKIGKPIITGALAPLAEFALHTPILGSEGLGGSVSAGIGRAGLVGKLS 331

QY 320 VPQAWAANQAVTPAARALPLTSLTS---AAERGPQMGGLPVGOMGARAGGSLGVLR 376
DB 332 VPQGWTVAAPEIPSPAAALQATRLAAPIAATDAGALLGGMALSLAGRAAAGSTG--- 388

QY 377 VPPRPYVMPHSPAAAG 391
DB 389 ---HPIGSAAAPAVG 400

```

Search completed: August 25, 2004, 03:13:02
Job time : 121 secs

```

RESULT 15
Q7TZ14 ID Q7TZ14 PRELIMINARY; PRT; 423 AA.
AC Q7TZ14;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PPE family protein.
GN PPE29 OR MB1829.

```

```

OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Biglieri K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Akin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248340; CAD94532.1; -.
DR Complete proteome.
SQ SEQUENCE 423 AA; 41477 MW; 26E52CC271FBBF57 CRC64;

```

```

Query Match 37.8%; Score 736; DB 16; Length 423;
Best Local Similarity 41.8%; Pred. No. 6.2e-34;
Matches 182; Conservative 50; Mismatches 123; Indels 80; Gaps 12;

QY 2 VDFGALPPEINARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGWSIGS 61
DB 1 MDFGLLPPEINSGRMVTPGPGPMLAAATAWDGLAVELHATAAGYASELSALT-GAWSGP 59

QY 62 SAGLMVAASPYVAVMSVTAGQAELETAQVRAAAVETAYGLTVPPVIAENRAELMIL 121
DB 60 SSTMSASAAAPYVAVMSATVHAELAGAQARLAIAAYEAFAATVPPVIAANRAQLMVL 119

QY 122 IATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGGL 181
DB 120 IATNIFGQNTPAIWMTEAQYEMWMAQDAAMFYAGSSATA-SRMTAFTEPPQTTHGQL 178

QY 182 LEQAAVEEASDTAAAN-----QLMNVVPAQLQLOAQPT-----QGTFP-----S 221
DB 179 GAQSSAVAQTAATAAGNLOSAPFQLLSAVPRALQGLALPTASQSASATPQWVTDLGNLS 238

QY 222 SKLGLMKTVPSPHRSPTSNMVMNMHMTNSGVSMTLSSMLKGFAPAAAAAQVTA 281
DB 239 TFLGG--AVTGPYTFP-----GVLPPSGVPYLLGIQSVL-----V 271

QY 282 AQNGVRAMSSILGS-----SLGSSGLGGG--VAANLGRAASVGSLS 319
DB 272 TQNGQGVSAALLGKIGKPIITGALAPLAEFALHTPILGSEGLGGSVSAGIGRAGLVGKLS 331

QY 320 VPQAWAANQAVTPAARALPLTSLTS---AAERGPQMGGLPVGOMGARAGGSLGVLR 376
DB 332 VPQGWTVAAPEIPSPAAALQATRLAAPIAATDAGALLGGMALSLAGRAAAGSTG--- 388

QY 377 VPPRPYVMPHSPAAAG 391
DB 389 ---HPIGSAAAPAVG 400

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:57:59 ; Search time 25 Seconds
(without alignments)
814.377 Million cell updates/sec

Title: US-09-724-685-107
Perfect score: 1949
Sequence: 1 MVDGALPEINSARMYAGP.....SGVLRVPPRYMHPSPAAG 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1556.5	85.0	336	1 YD61_MYCTU	Q1031 mycobacteri
2	705	36.2	408	1 YS92_MYCTU	Q10813 mycobacteri
3	688.5	35.3	463	1 Y102_MYCTU	O53951 mycobacteri
4	444	22.8	487	1 Y442_MYCTU	P42611 mycobacteri
5	426.5	21.9	443	1 Y878_MYCTU	Q10540 mycobacteri
6	418	21.4	678	1 YF48_MYCTU	Q10778 mycobacteri
7	364.5	18.7	408	1 SRA_MYCLE	Q07297 mycobacteri
8	324.5	16.6	463	1 Y096_MYCTU	Q10892 mycobacteri
9	324	16.6	434	1 YU18_MYCTU	P31500 mycobacteri
10	321.5	16.5	435	1 YU21_MYCTU	O53268 mycobacteri
11	233.5	12.0	178	1 YX29_MYCTU	O06246 mycobacteri
12	217.5	11.2	176	1 YX25_MYCTU	O50703 mycobacteri
13	153.5	7.9	860	1 ELS_MOUSE	P54320 mus musculu
14	151.5	7.8	232	1 YX23_MYCTU	Q50702 mycobacteri
15	150	7.7	884	1 ELS_RAT	Q99372 rattus norv
16	143.5	7.4	730	1 ELS_HUMAN	P15502 homo sapien
17	139.5	7.2	881	1 PRY3_YEAST	P47033 saccharomyc
18	135	6.9	907	1 A180_HUMAN	O60641 homo sapien
19	133.5	6.8	825	1 ICP0_HSV2H	P28284 herpes simp
20	133.5	6.8	2432	1 Y43R_IRV6	P18305 chilo iride
21	132	6.8	1120	1 STFR_ECOLI	P78072 escherichia
22	132	6.8	2090	1 N214_HUMAN	P35558 homo sapien
23	131	6.7	836	1 VQ26_BPM15	Q05233 mycobacteri
24	131	6.7	1150	1 APMU_PIG	P12021 sus scrofa
25	129	6.6	1783	1 RAA3_CHLRE	O9fec4 chlamydomon
26	128	6.5	779	1 SRP_DROME	P52172 drosophila
27	127.5	6.5	790	1 ANP_NOTCO	P24856 notothenia
28	126	6.5	354	1 YAUG_SCHPO	Q10189 schizosacch
29	125.5	6.4	1211	1 BUN2_DROME	Q24523 drosophila
30	125	6.4	577	1 CST2_HUMAN	P33240 homo sapien
31	124	6.4	2090	1 HFC1_MESAU	P51611 mesocricetu
32	123.5	6.3	677	1 Y136_MYCTU	O50597 mycobacteri
33	123	6.3	432	1 YF10_MYCTU	P71789 mycobacteri

34	123	6.3	1140	1 YN96_YEAST	Q04893 saccharomyc
35	122	6.3	394	1 HYF1_ALCEU	P45805 aicaligenes
36	121.5	6.2	1025	1 SLAP_CAUCR	P35828 caulobacter
37	121.5	6.2	1845	1 Z236_HUMAN	Q09136 homo sapien
38	121	6.2	1199	1 P121_RAT	P52591 rattus norv
39	119.5	6.1	635	1 HMLA_DROME	P10105 drosophila
40	119	6.1	915	1 A180_RAT	O05140 rattus norv
41	118.5	6.1	580	1 EXPR_XANCP	P23314 xanthomonas
42	118.5	6.1	2038	1 FSH_DROME	P31709 drosophila
43	118	6.1	1508	1 BCSX_XANAC	P88938 xanthomonas
44	117	6.0	444	1 Y808_CHLPN	Q02798 chlamydia p
45	117	6.0	652	1 PICA_HUMAN	Q13492 homo sapien

ALIGNMENTS

RESULT 1
YD61_MYCTU
ID YD61_MYCTU STANDARD; PRT; 396 AA.
AC Q11031;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein RV1361C/MT1406.
GN RV1361C OR MT1406 OR MTCY02B10.25C.
OS Mycobacterium tuberculosis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bisshai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z75555; CA99966.1; --
CC EMBL: AEC07013; AAK45669.1; --
CC PIR: H70741; H70741.
CC TIGR: MT1406; --
CC TubercuList; RV1361C; --
CC InterPro: IPR000030; Microbac_PPE.

```

DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 158 159 TA -> AT (IN REF. 2).
SQ SEQUENCE 396 AA; 40015 MW; 6AFAE0D7B5F668D0 CRC64;

Query Match 85.0%; Score 1656.5; DB 1; Length 396;
Best Local Similarity 85.1%; Pred. No. 1.3e-89;
Matches 338; Conservative 18; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
DB 1 MVDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
QY 61 SSAGLMVAASPYVAVMSVTAGQAEALTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
DB 61 SSAGLMVAASPYVAVMSVTAGQAEALTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLGQNTPATAVNEAEYGEWMAQDAAMGYAAATATATATLTPPEAEPEMTSAGG 180
DB 121 LIATNLGQNTPATAVNEAEYGEWMAQDAAMGYAAATATATATLTPPEAEPEMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNVPQALQQAQPTQCTTPSSKLGGLWKTVPSPHRSPIGN 240
DB 181 LLEQAAVEEASDTAAANQLMNVPQALQQAQPTQCTTPSSKLGGLWKTVPSPHRSPIGN 240
QY 181 LLEQAAVEEASDTAAANQLMNVPQALQQAQPTQCTTPSSKLGGLWKTVPSPHRSPIGN 240
DB 181 LLEQAAVEEASDTAAANQLMNVPQALQQAQPTQCTTPSSKLGGLWKTVPSPHRSPIGN 240
QY 241 MVSNNHMTNSGVMTWLTSLMLKGFAPAPAAAQVQTAQNGVRAMSS-----LGSSL 296
DB 241 IVSNLNNHMTNSGVMTWLTSLMLKGFAP-AAAQVQVETAAQNGVQAMSSLGSLGSSL 299
QY 297 GSSGLGGVGAANLGRASVGSLSVPQAWAANAQVTPAARALPLTSLTAAERGPQMLG 356
DB 300 GSSGLGAGVAANLGRASVGSLSVPQAWAANAQVTPAARALPLTSLTAAERGPQMLG 359
QY 357 GLPVQGMARAG--GGLSGVLRVPPRYVMPHSPAAG 391
DB 360 GLFLGQLTSGGGGGSVNALRMPPRAYVMPRVPFAAG 396

RESULT 2
YS92_MYCTU STANDARD; PRT; 408 AA.
AC Q10813;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein RV2892c/MT2959/Wb2916c.
GN RV2892C OR MT2959 OR MTCY274.23C OR MB2916C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
[1]
SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=8295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Harnaby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,

```

```

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12768972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- SIMILARITY: Belongs to the mycobacterial pPE family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; Z74024; CAA98377.1; -.
CC EMBL; AB007119; AAK7285.1; -.
CC EMBL; BX248344; CAD96603.1; -.
CC PIR; G70925; G70925.
CC TIGR; MT2959; -.
CC TubercuList; RV2892c; -.
CC InterPro; IPR000030; Microbac_PPE.
CC Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 56 76 POTENTIAL.
SQ SEQUENCE 408 AA; 41469 MW; 3E3D1F20D7827199 CRC64;

Query Match 36.2%; Score 705; DB 1; Length 408;
Best Local Similarity 41.8%; Pred. No. 2.7e-34;
Matches 182; Conservative 47; Mismatches 134; Indels 72; Gaps 14;

QY 2 VDFGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIGS 61
DB 1 MDFGVLPPEINSARMYAGPGSGPMWAAAADWSLAEGLAAGYRLAISLTGAYWAGP 60
QY 62 SAGLMVAASPYVAVMSVTAGQAEALTAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 121
DB 61 AAASWAAVTPYVAVLSATAGQAEQAGQARAAAAYELAFAMTVPVAVANRALVAL 120
QY 122 IATNLGQNTPATAVNEAEYGEWMAQDAAMGYAAATATATLTPPEAEPEMTSAGGL 181
DB 121 VATNFFGQNTPATATATQVAAEMWQAQDAAMAYAGSAAT-ELTFTTAAPTTSPAAL 179
QY 182 LEQAAA-----VEEASDTAAANQLMN-VP--QALQQ-LAQPTQGTTPSSKL--- 224
DB 180 AGQAAATVSVTPVPLATTAAVQQLQLLSLTPWYSALQWLAENLLGLTPDNRTIV 239
QY 225 -----GGLWKTVPSPHRSPIVSNMANNHMTNSGVMTWLTSLMLKGFAPAAA 275
DB 240 RLIGISVDEGL-----LQFEASLAQAQIPGTPGGAG--DSSSVLDWSGPTIFA 287
QY 276 QAVQTAQNGVRAMSSL--GSSLSG-----SLGCGVAANLGRASVGSLS 319
DB 288 -----GPRASPVAGGAVGVQTPQPYWYWDRESIGSVSAALCKGSSAGSL 338
QY 320 VPQAWAANAQVTPAARALP-----LTSLSAAERGPQMLGGLPVQGMARAGGSLSVLR 376
DB 339 VPPDWAARAWANPAARLPGDDVTLARGTAENA---LLRGFPMAAGQSTGGF--VHK 393
QY 377 VPPRYVMPHSPAAG 391
DB 394 YGFLAVMQRPFFAG 408

```

RESULT 3
 Y102 MYCTU STANDARD; PRT; 463 AA.
 AC O53951;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical PPE-family protein RV1851/MT1851/MD1830.
 GN RV1802 OR MT1851 OR MT049.24 OR MB1830.
 OS Mycobacterium tuberculosis, and
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773, 1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekai F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.bovis; STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RA "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
 CC -I- SIMILARITY: Belongs to the mycobacterial PPE family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC ENBL; AL022021; CAAL1723.1;
 DR ENBL; AEO07044; AAK46123.1;
 DR ENBL; BX248340; CAD94533.1;
 DR PIR; C70931; C70931.
 DR TIGR; MT1851;
 DR TubercuList; RV1802;
 DR InterPro; IPR000030; Microbac_PPE.
 DR Pfam; PF00823; PPE; 1.
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 401 401 S -> L (IN REF. 2).

SQ SEQUENCE 463 AA; 46021 MW; EE64828BF09FA551 CRC64;
 Query Match 35.3%; Score 688.5; DB 1; Length 463;
 Best Local Similarity 42.8%; Pred. No. 2.8e-33;
 Matches 174; Conservative 48; Mismatches 136; Indels 49; Gaps 11;
 QY 2 VDFGALPPEINSAARMYAGPGSASLVAAAQMWDSVASDFSAASAFQSVWVWGLTVGWSIGS 61
 DB 1 MDFGVLPPEINSGRMVYAGPGSGPMLAAAAAWDGLATELQSTADYGVISVLT-GVWSGQ 59
 QY 62 SAGLMVAAAASPYVAMVVTAGQBELTAAQVRVAAAAAYETAYGLTVPPPPVIAENPAELMIL 121
 DB 60 SSGTMAAAAAPYVAMVGTAAALAREAAAQASAAAAYEAAFAATVPPPPVVAANRAELAVL 119
 QY 122 IATNLGQNTPAIAVNEAEYGEVMAQDAAMFGVAAAATATATATATLLFFEPAPMTSAGGL 181
 DB 120 AATNIFGQNTGAIAAAEARYAEVMAQDAAMFGVAAAATATATATATLLFFEPAPMTTNAAGL 178
 QY 182 LEQAAVVEEASDTAAANQNMNVPAQLQQAQPTGTTTSSKLGGLMKTYS--PHRSP1- 238
 DB 179 ATQGVAVAAQAVGASAGN-ARSLVSEVLEFLA--TAGTNYNKTVASLNMNAVTVGVYASSVY 235
 QY 239 -----SNMYSMANNHSMNTNSGVSMNTLSSMLKGPAPAAAQAQVTAQAQNVRA 288
 DB 236 NSMLGLGPAESKVMVLPANDTVISTIFGVQVQKFFNPVTFPNDLIFK----- 283
 QY 289 MSSLGSSSLG-----SSGLGG---GVAANLGEAASVGSLSVPQAAANQAVTPAARALPL 340
 DB 284 -SALGAGLGLSASLSSGLGSTAPASAGASQAGSVGMSVPPSWAAATPAIRTVAAVFSS 342
 QY 341 TSLTS--AAERGPQOML-----GGLPVQMGARAGGGLSGVLRV 377
 DB 343 TGLQAVPAAAISEGLLSQMALASVAGGALGAAARATGGFLGGRV 389

RESULT 4
 Y442 MYCTU STANDARD; PRT; 487 AA.
 ID Y442 MYCTU
 AC P42611; O53727;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical PPE-family protein RV0442c/MT0458.
 GN RV0442C OR MT0458 OR MV037.06C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Erdmann;
 RX MEDLINE=87137260; PubMed=3029018;
 RA Shinnick T.M.;
 RA "The 65-kilodalton antigen of Mycobacterium tuberculosis."
 RL J. Bacteriol. 169:1080-1088 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekai F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;

```

RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15467; AAA89235.1; ALT INIT.
DR EMBL; AL021932; CA817399.1; -.
DR EMBL; AE006948; AAK44681.1; -.
DR EMBL; AE006948; AAK44681.1; -.
DR TIGR; MT0458; -.
DR TubercuList; RV0442c; -.
DR InterPro; IPR000030; Microbac_PPE.
DR InterPro; IPR002989; Mycobac_Pentapep.
DR Pfam; PF01469; Pentapeptide_2; 5.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 40 40 E -> K (IN REF. 2).
FT CONFLICT 96 96 I -> T (IN REF. 1).
FT CONFLICT 211 211 G -> GNNNG (IN REF. 1).
SQ SEQUENCE 487 AA; 47247 MW; 97234D5B316C8C7F CRC64;

Query Match 22.88; Score 444; DB 1; Length 487;
Best Local Similarity 32.18; Pred. No. 4.8e-19; Indels 58; Gaps 12;
Matches 135; Conservative 48; Mismatches 180;

QY 4 FGALPPEINSARMYAGSGSASLVAAQMDVSDLFSAASAFQSVVWGLTVGSGWIGSSA 63
DB 6 FAWLPPEINSALMFAGSGSLPIAATAWAGELAEELLASIASIGSVTSELTSGLMIGPSA 65
QY 64 GLMVAASPVVAMSVTAGQAEITAAQVRAAAYETAGLVTPPVPIAENRAELMILIA 123
DB 66 AAMWAVATQYLAWLSTAAQAQQAQAAMAIAFAEAALAAATVQPAVVAANGLMQLAA 125
QY 124 TNLGQNTPTAIAVNEAEYGEWMAQDAAMFGVAAATATATATLLPPEAPEMTSAG--- 179
DB 126 TNWFGQNPALMDVEAAAYEQMWALDVAAAGYHFDASAAVAQLAPQOV--LRNLGIDIG 183
QY 180 --GLLEQAAVEASDTAAANQMLNNVPOALQOLAQPTGTTPTSSKLG-----GLW 228
DB 184 KNGQINLFGNGTGNIGNINNIGNNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTG 243
QY 229 KTVS-----PHRSPISNMVSMANNHMTNSGVSTNTLSMLKGFAPAAAQAQV 279
DB 244 NTGSGNIGFQITGDHQVGFNGSGSGN-IGFNGSGTGNVGFNS----- 287
QY 280 TAAQNGVRAVSSLGSGSLGGVAAANLGRAASVGSLSVPAQAAANQAVTPAARALP 339
DB 288 GSGNIGTGNIGSGSLNSGIGTSGT---INAGLG---SAGSLNT-SFWNAGNQAALGSAAGS 340
QY 340 LLSLTSAABERGPGM-----LGLPGVQMGARAG--GGLSGVLR--VPPRPVMPHSPAA 390
DB 341 EAAVLSAGVATGNTAALSSGILLASALGSTGLGQHLANVLNGLTNPVAPASAPV 400
QY 391 G 391
DB 401 G 401

```

RESULT 5

```

ID Y878_MYCTU STANDARD; PRT; 443 AA.
AC Q10540;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein RV0878c/MT0901.
GN RV0878C OR MT0901 OR MTCY31.06C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
[1]_
RN RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73101; CAA97385.1; -.
DR EMBL; AE006977; AAK45143.1; ALT_INIT.
DR TIGR; MT0901; -.
DR TubercuList; RV0878c; -.
DR InterPro; IPR000030; Microbac_PPE.
DR InterPro; IPR002989; Mycobac_Pentapep.
DR Pfam; PF01469; Pentapeptide_2; 4.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Repeat; Complete proteome.
FT CONFLICT 15 35 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT DOMAIN 64 73 POLY-ALA.
FT DOMAIN 81 115 ALA-RICH.
FT DOMAIN 231 270 4 X 10 AA APPROXIMATE REPEATS.
SQ SEQUENCE 443 AA; 43592 MW; C58BEC07F0675E2 CRC64;

Query Match 21.98; Score 426.5; DB 1; Length 443;
Best Local Similarity 31.28; Pred. No. 4.5e-18;
Matches 125; Conservative 56; Mismatches 132; Indels 87; Gaps 14;

```

```

Qy      2  VDFGALPBEINSRMYAGPGSASLVAAQAQMWDSVASDLFSAASAFOSVWGLTVG--SW 58
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      1  MNFWLPEVNSARIYACGAPAPMLAAVAADWGLAAELGMAAASFLLISGLTAPGSAW 60
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy      59  IGSSAGLWVAASPVVAMSVTACQAELTAAQVPRVAAAAYETAYGLTVPPVIAENRAEL 118
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      61  QGPAAAMAAAAAPYLSWLNAATARAAGAAAGAKAAAVYERAAATAPALVAANRNOL 120
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy      119  MILIATNLLGONTPAIAVNEAEYEGEMWAQDAAMFGYAAAATATATATLTPFBEAPEMISA 178
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      121  LSLVLSNLFQOMLPALATEASYEQLWAQDVAAVMVGHGASIVASQLTPWQQ----- 173
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy      179  GGLLEQAAAEEASDTAAANOLMNNVPOALQOLAQTOGTTPSKLGGWKVTSVPHRSPI 238
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      174  --LLSVLPVTVTAAPAGAV---GVPA--LAIPALGV---ENIG-----V 208
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy      239  SNWV---SWANDHMKWTHSG-----VSMTN-----TLSSMLXGFAPALAAQAV 278
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      209  GNFLGIGNIGNNVGSGNTGPDYFGIGNIGNWANTGNIGNALNGSGNAGFFNFGNGNDG 268
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy      279  QTAQONGVEAMSSLGS--SLGSSGLGGVAA--NLGRAASVGSLSVPQAAWAAANQAVTPAA 335
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      269  NTNFGSGNAGFLNIGSGNEGSGNLGFGNAGDDNTG-----WGNSGD----- 309
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy      336  RALPLTSLTSAABERPGQMLGGL--PVGQMGARAGGGLSG 373
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      310  -----NTGGENSGDLNTGIGSPVTOGVANSGFGNTG 341
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

RESULT 6	
YF48_MYCTU	
ID_YF48	STANDARD; PRT; 678 AA.
AC	Q10778;
DT	01-OCT-1996 (Rel. 34, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Hypothetical PE-family protein RV1548C/MT1599.
GN	RV1548C OR MT1599 OR MTCV48.17.
OS	Mycobacterium tuberculosis.
OC	Bacteria; Actinobacterii; Actinobacteridae; Actinomycetia
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX	NCBI Taxid=1773;

[1] -
RN SEQUENCE FROM N.A.
RP STRAIN=H37Rv;
RC
RX MEDLINE=96299987; PubMed=9634230;
RA Colle S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagals K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruster S., Seeger K., Skelton S., Squares R., Squares R.,
RA Sutton J.E., Taylor K., Whitehead S., Barrrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland R., Eisen J.A., Carpenter L., White C.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3] - SIMILARITY: Belongs to the mycobacterial ppe family.

CC
CC-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib.ch/announcement>)
CC or send an email to license@ebi-sib.ch.

CC	-----
DR	EWBL; Z74020; CAA93335.1; --
DR	EWBL; AE007026; AAK45866.1; ALT_INIT.
DR	PIR; A70762; A70762.
DR	TIGR; MT1599; --
DR	TubercuList; Rv1548c; --
DR	InterPro; IPR00030; Microbac_PPE.
DR	InterPro; IPR002989; Mycobac_pentapep.
DR	Pfam; PF01469; Pentapeptide_2; 11.
DR	Pfam; PF00823; PPE; 1.
DR	Hypothetical protein; Transmembrane; Complete proteome.
KW	
FT	TRANSMEM 14 34 POTENTIAL.
FT	TRANSMEM 180 200 POTENTIAL.
FT	CONFLICT 258 258 .D -> G (IN REF. 2).
FT	CONFLICT 258 AA: 667336 MW: 209F1593D52533A2 CRC64;
FO	SEQUENCE 678

Query Match	21.4%;	Score 418;	DB 1;	Length 678;
Best Local Similarity	31.9%;	Pred. No. 2.2e-17;		
Matches 106;	Conservative 48;	Mismatches 140;	Indels 38;	Gaps 7;

QY	2	VDFGALPPEINASRYAGPGSASIVAAQMWDSVASDLFSAASAFQSVWGLTIVGCSWIGS	61
DB	1	MNFVLPPEINSLMFAGAGPGLMAAASAWTGLAGDLGSAASFSAVTSQLATGSGWGP	60
QY	62	SAGLMVAASAPYVAWMSVTGAQELTAAQVRVAAAAYETAYGLTVPPPIAENRAELML	121
DB	61	ASAAWTGVAASYARWLTTAAQAQEAQAAQAAQAAVSFAFEALAAATHPGAVASANRGLRSL	120
QY	122	IATNLLGNTPAIAVNEAEYGEWMAQDAAMFGVAAATATATATATLLPPEEAPEMTSAG-	180
DB	121	VASNLLGONAPAIIAAVEAVIEQMWAAADVAAMLGTHGEASVALSLTPTTPFSAAATPGG	180
QY	181	-----LLEQAAAEVEEASPTAAANQLMNVNPOALQQAQPTGCTTPSSKLGLGWKTVS	232
DB	181	AVIAGPPFLDLGNVTIGGFNLASNLGLNLGS-----FNPGSANTGSVNLGN-----	229
QY	233	PHRSPISNK--VSMANNHMSMTNSGVSMNTLTSSMLKGFAPAAAAQAQVTAQAONGVRAMS	290
DB	230	-----ANIGDLNLSGNGIGSYNLGGNTGDLN-----PDSGNTGTILNWSGNGIGSYN	276
QY	291	SLGSSLGSSGLGGGVA--ANLGRAASVGSLSV	320
DB	277	IGGGNIGSYNLGSGNTGDTNFG-GGNTGNLNV	307

RESULT 7

SRA_MYCLE	STANDARD;	PRT;	408 AA.
ID	SRA_MYCLE		
AC	Q07297;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Serine-rich antigen (25L) (45 kDa protein).		
GN	SRA OR ML0411 OR MLCL383.14.		
OS	Mycobacterium leprae.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.		
OX	NCBI_TaxID=1759;		
RN	[1]_		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9329328; PubMed=8478104;		
RA	Vega-Lopez F., Brooks L.A., Dockrell H.M., de Smet K.A.L.,		
RA	Thompson J.K., Hussain R., Stoker N.G.;		
RT	"Sequence and immunological characterization of a serine-rich antigen		
RT	from Mycobacterium leprae."		
RL	Infect. Immun. 61:2145-2153(1993).		
RP	[2]		
RN	SEQUENCE FROM N.A.		

RA MEDLINE=95020554; PubMed=7934845;
 RA Rinke de Wit T.F., Clark-Curtiss J.E., Abebe F., Kolkon A.H.J.,
 RA Jonson A.A.M., Rhole J.B.R.;
 RA "A Mycobacterium leprae-specific gene encoding an immunologically
 RT recognized 45 kDa protein."
 RL Mol. Microbiol. 10:829-838(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN.
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hanlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus."
 RL Nature 403:1007-1011(2001).
 CC -1- SIMILARITY: Belongs to the mycobacterial PPE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL; U00015; AAC3220.1; -
 DR EMBL; X68431; CAA48480.1; -
 DR EMBL; Z21952; CAA79950.1; -
 DR EMBL; Z97179; CAB09938.1; -
 DR EMBL; AL583918; CAC29919.1; -
 DR PIR; C86960; C86960.
 DR PIR; S33522; S33522.
 DR PIR; S39872; S39872.
 DR Lepnora; ML0411; -
 DR InterPro; IPR000030; Microbac_PPE.
 DR Pfam; PF00823; PPE; 1.
 KW Antigen; Repeat; Complete proteome.
 FT DOMAIN 192 196 POLY-SER.
 FT DOMAIN 209 235 2 X 6 AA REPEATS OF S-V-A-Q-S-E.
 FT REPEAT 209 214 1.
 FT REPEAT 230 235 2.
 FT CONFLICT 132 132 T -> S (IN REF. 2).
 FT CONFLICT 189 189 S -> L (IN REF. 2).
 FT CONFLICT 191 191 H -> D (IN REF. 2).
 FT CONFLICT 292 292 F -> L (IN REF. 2).
 SQ SEQUENCE 408 AA; 42466 MW; 5C0C2B80D656A9D8 CRC64;

Query Match - 18.7%; Score 364.5; DB 1; Length 408;
 Best Local Similarity 26.3%; Pred. No. 1.7e-14;
 Matches 115; Conservative 72; Mismatches 163; Indels 87; Gaps 11;
 QY 1 MVDFFGALPPEINARMYAGFSGSLVAAQMDSDVASDLSFSAAPQSVVGLTVGWSWG 60
 DB 1 MFDFFVYSPEVNAFLMSRGPGSTPLNGAAEAEMISLAEQLEMAAQEVSPTIVVAVPASFAG 60
 QY 61 SSAGLMVAAPSPVWVMSVTAGAELETAQVRAAAVETAYGLTVPPPVIAENAEELMI 120
 DB 61 ETSDMLASRVSIFVNLGDGNAENAGLIARVLHVAVAFEAARAGVPLTLVLGNIIHTWA 120
 QY 121 LIATNLLGNTPAIAVNEAEYGEHMAQDAAMFGVAAATATATATALLPPEEAPMTSAGG 180
 DB 121 LKAINFVGQSTTVAALEADYDILMWYQNSTMTTVDTLVRETGRMENFEPAPQLVSR-Y 179
 QY 181 LLEQAAVEAEASDTAAANGLNNVPQALQQLAQ - - - - - 213
 DB 180 CWDRRDSVNSFHSSSSDSLYESIDNLYDSVAQSEHSDMSQSYNTCGSVAQSELCD 239

QY 214 ----PTQ-----GTPSSKLGGLWKTVPSPHRSPISNMVMNMHSMNTNSGVSMNT 261
 DB 240 PFGTSPQSQSSNDLSATSLTQQLGL-----DSIISSASASALLTNS--ISSST 286
 QY 262 LSSMLKGFAPAAAQAQVTAQAQNGVRAMSSSLGSSGLGGVAAN-----LGRAASVG 316
 DB 287 ASSIM-----PIVASQVTEITLGRSQV-AVEKMIQSISSSTAVSDVAASKVAGVQAVSVG 341
 QY 317 SLUSVQAAANQAVTPAARALP--LTSLSAAERGPQGMGLGFLPVGQMGARAGGGLSGV 374
 DB 342 ALRUPENWATASQPMWATAHSVPAGCSAITTA-----VSGPLEGV 381
 QY 375 LRVPPRPVPMHSPAAG 391
 DB 382 TO--PAEEVLITASVAGG 396
 RESULT 8
 Y096 MYCTU STANDARD; PRT; 463 AA.
 ID Y096 MYCTU STANDARD; PRT; 463 AA.
 AC Q10832;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical PPE-family protein RV0096/MT0105.
 GN RV0096 OR MT0105 OR MTCV251.15.
 OS Mycobacterium tuberculosis
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 ON NCBI_TaxID=1773;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22208494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the mycobacterial PPE family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Z74410; CAA98932.1; -
 DR EMBL; AE006922; AAK44327.1; -
 DR PIR; H70750; H70750.
 DR TIGR; MT0105; -
 DR Tuberculist; RV0096; -

```

DR InterPro: IPR000030; Microbac_PPE.
DR Pfam: PF00823; PPE; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 3 23
FT TRANSMEM 88 108
FT TRANSMEM 122 132
FT TRANSMEM 216 236
FT TRANSMEM 245 265
FT TRANSMEM 276 296
FT TRANSMEM 323 343
FT TRANSMEM 419 439
FT TRANSMEM 463 484
SQ SEQUENCE 463 AA; 46894 MW; 42D9D66A033D0DD8 CRC64;

Query Match
Best Local Similarity 16.6%; Score 324.5; DB 1; Length 463;
Matches 112; Conservative 54; Mismatches 144; Indels 93; Gaps 12;

QY 6 ALPPEINARMYAGPGSGSLVAAQWDSVADLFSAASAFQSVVWGLTVGSGSAGL 65
Db 2 AIPPEVHSGLLGAGCGPGSLVAAQWDSVADLFSAASAFQSVVWGLTVGSGSAGL 61
QY 66 MYAAASPYVAMSVTGAQAELETAQVRAAAYETAYGLTVPPVIAENRAELMILIATN 125
Db 62 YVAAHGPLYLAWLEQTAINSAVTAACHVAAAYCSALAAAMPTPABLAANHAITHGLVLIATN 121
QY 126 LLGQNTPAIVNEAYGEMWQDAAMFGYAAATATATATATLLPFEAPEMTSAGLLEQA 185
Db 122 FFGINTVPTALNEADYRVRLQADTMAYQAVADATVAVSTQAPPIRAPGG----- 176
QY 186 AAVEEASDT-----AAANGLMNVPAQLQAQPTQ-----GTPSSKL----- 224
Db 177 -----DAADTDLVSSIGQLIRDI---LDFIANPKYFLEFFEQFGFSPAVTWLALNAL 229
QY 225 ---GGW---KTVSPHRSPISNMVMNMTNSGVSMTNTLSMLK---GF 269
Db 230 QLYDFLWYFYASYGILLPLFPFT-----TLTALUSALIHILNLPAGL 275
QY 270 APAAAQAVQTAQNGVRAMSSSLGSLGGVAAANLGRAASVGSLSVFWAAANQ 329
Db 276 LPIAAA-----LPGDQGANLAVATPATAVP-----GGSP 308
QY 330 AVTPAARALPLTSLTSAERPG---QMLGGLPVG-QMGARAG 368
Db 309 PTNPAPAAPSSNSVGSASAAPGISYAVFLAPGVSSGPKAG 351

RESULT 9
YU18_MYCTU STANDARD; PRT; 434 AA.
AC P31500; O53265;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical PPE-family protein Rv3018c/MT3098/MT3101.
GN Rv3018c OR MT3098/MT3101 OR MV012.32C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K., Davlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
```

```

RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RC MEDLINE=2206494; PubMed=12218036;
RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RX Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Emolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP SEQUENCE OF 160-374 FROM N.A.
RC STRAIN=Isolate 50410;
RL Pakki A.H., Dale J.W.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC -!- CAUTION: In strain Oshkosh the gene for this protein is
CC interrupted in position 307 by an IS6110 element.
CC -!- CAUTION: Was originally (Ref.3) thought to be a dihydrofolate
CC reductase.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC in positions 294; 337 and 355.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AL021287; CA416103.1; -
CC EMBL; AE007129; AAK47427.1; ALT_SEQ.
CC EMBL; AE007129; AAK47430.1; ALT_SEQ.
CC EMBL; X59271; CA441961.1; ALT_FRAME.
CC PIR; E70857; E70857.
CC TIGR; MT3098; -.
CC TIGR; MT3101; -.
CC TubercuList; Rv3018c; -.
CC InterPro; IPR000030; Microbac_PPE.
CC Pfam; PF00823; PPE; 1.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 434 AA; 43029 MW; 41D673C4BD389DD6 CRC64;

Query Match
Best Local Similarity 16.6%; Score 324; DB 1; Length 434;
Matches 115; Conservative 56; Mismatches 178; Indels 62; Gaps 12;

QY 6 ALPPEINARMYAGPGSGSLVAAQWDSVADLFSAASAFQSVVWGLTVGSGSAGL 65
Db 8 ASPPEVHSGALLSAGPGSLQAQAAAGWSALSAAVAAQELSVVVAAGVAGVWQPSAEL 67
QY 66 MYAAASPYVAMSVTGAQAELETAQVRAAAYETAYGLTVPPVIAENRAELMILIATN 125
Db 68 FVAAVVPVYVAMLVQASADSAASAAAGHEAAAGVVCALAEEMPTLPBLAANHLTHAVLATN 127
QY 126 LLGQNTPAIVNEAYGEMWQDAAMFGYAAATATATATLLPFEAPEMTSAGLLEQA 185
Db 128 FFGINTVPTALNEADYRVRLQADTMAYQAVADATVAVSTQAPPIRAPGG----- 181
QY 186 AAVEEASDTAAAN-----OLMNVPAQLQAQPTQCTTPSSKLGGLWKTSPH 234
Db 182 --ANEASNAVAATATTPFPWHIVQFLBETFAAYQOYLSALLSELPA--VAVWVQLFVD 237
QY 235 ---RSPISNMVMNMTNSGVSMTNTLSMLKGF----- 270
Db 238 ILGFGNIIGFIITLASNAQLLTETFAINASVAVVGLLYALAGVIDIVVWVIGNLFGVPL 297
QY 271 -----PAAAQAVQTAQNGVRAMSSSLGSLGGVAAANLGRAASV-GSLSVPOA 323
```

Db 298 GGPLGALAAAVPGVAGLAGVAGLAAL-PAVGAA--AGAPAAALVGSVAPVSGGWSPQA 354
 QY 324 WAAANQAVTPAARALPLTSLTSAERPGQMLGGLPVGQMGARAGGGLSGV 374
 Db 355 RLVS--AVEPAPASTSVSL--ASDRGAGAL--GF-VGTAGKESVGQAPAGL 398

RESULT 10
 YU21 MYCTU
 ID YU21 MYCTU STANDARD; PRT; 435 AA.
 AC 053269; 053269;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical PPE-family protein RV3021C/RV3022C/MT3106.
 GN RV3021C/RV3022C OR MT3106 OR MTV012.35C/MTV012.36C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 82.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL021287; CAA16106.1; ALT_FRAME.
 CC EMBL; AL021287; CAA16107.1; ALT_FRAME.
 CC EMBL; AS007129; AAK47435.1;
 CC TIGR; MT3106;
 CC TubercuList; RV3021C;
 CC TubercuList; RV3022C;
 CC InterPro; IPR000030; Microbac_PPE.
 CC Pfam; PF0823; PPE: 1.
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 299 G -> A (IN REF. 2).
 FT CONFLICT 317 320 LAGV -> VTGL (IN REF. 2).
 FT CONFLICT 326 326 L -> V (IN REF. 2).
 SQ SEQUENCE 435 AA; 42876 MW; 3B157643EAA8484A CRC64;

Query Match 16.5%; Score 321.5; DB 1; Length 435;
 Best Local Similarity 26.5%; Pred. No. 5.6e-12;
 Matches 115; Conservative 53; Mismatches 159; Indels 107; Gaps 12;

QY 6 ALPPEINSARMYAGPGCSASLVAAQWDSVADSLFSAASAFQSVVWGLTVGSSWIGSSAGL 65
 Db 8 ASPPEVHSALLSAGPGCSLQAAAAAGWSALSAAEYAAVAQELSVVVAAGVAGWQGPSAEL 67
 QY 66 MYAASPYVANMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPPVIAENRAELMILIATN 125
 Db 68 FVAAYVPYVAMLVQASADSAAAGHEHAAGYVYCALAEMPTLPALANHLTHAVLVATN 127
 QY 126 LLGQNTPAIAVNEAREYGEWQAQDAAMFGYAA----- 157
 Db 128 PFGINTIPALNEADYVEMWQAATVMSAYEAVGAALVATPHTGPAPVIVKPGANEASN 187
 QY 158 ATATATATALLPPE-----APEMTSAGGLLQAAAEASDTAAANQLMNVPOALQ 209
 Db 188 AVAAATITPPFPFGLAKFLEMAQAQAFTEVGELIMKSAEAWAVGFVELITGLVNEP--- 243
 QY 210 QLAQPTQGTTPSSKLGGLWKTVPSPHSFISNMVSK-----ANNH 248
 Db 244 -----WLIV-----LTGMDMFPATVGFALGVFLVPLEFAVVLE 278
 QY 249 MSMNNGVSMNTLTSS-----MLKGFAPAAAAQAVQTAQNGVRAMSSILGSSGSL 301
 Db 279 LAILSIGMIISNIFGAIPVLGGLLALAAAVVPGVAGLAGVAGLAALPAVGAAGAP-- 336
 QY 302 GGGVAANLGRAVSV-GSLSVPOAAANQAVTPAARALPLTSLTSAERPGQMLGGLPV 360
 Db 337 ----AALVGSVAPVSGGVSVSPQARLVS--AVEPAPASTSVSL--ASDRGAGAL--GF-V 385
 QY 361 GQMGARAGGGLSGV 374
 Db 386 GTAGKESVGQAPAGL 399

RESULT 11
 YU29 MYCTU
 ID YU29 MYCTU STANDARD; PRT; 178 AA.
 AC 006246;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical PPE-family protein RV3429/MT3533.
 GN RV3429 OR MT3533 OR MTCY77.01.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL J. Bacteriol. 184:5479-5490 (2002).
 CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 82.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL021287; CAA16106.1; ALT_FRAME.
 CC EMBL; AL021287; CAA16107.1; ALT_FRAME.
 CC EMBL; AS007129; AAK47435.1;
 CC TIGR; MT3106;
 CC TubercuList; RV3021C;
 CC TubercuList; RV3022C;
 CC InterPro; IPR000030; Microbac_PPE.
 CC Pfam; PF0823; PPE: 1.
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 299 G -> A (IN REF. 2).
 FT CONFLICT 317 320 LAGV -> VTGL (IN REF. 2).
 FT CONFLICT 326 326 L -> V (IN REF. 2).
 SQ SEQUENCE 435 AA; 42876 MW; 3B157643EAA8484A CRC64;

[illegible]

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of *Mycobacterium tuberculosis* from the

CC -1- SUBCELLULAR LOCATION: Extracellular matrix of elastic fibers.

CC -2- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=8;

CC Comment=Experimental confirmation may be lacking for some

CC isoforms;

CC Name=1; Sequence=Displayed;

CC IsoId=Q99372-1; Sequence=VSP_004245;

CC Name=2; Sequence=VSP_004244;

CC IsoId=Q99372-2; Sequence=VSP_004244;

CC Name=3; Sequence=VSP_004245;

CC IsoId=Q99372-3; Sequence=VSP_004245;

CC Name=4; Sequence=VSP_004246;

CC IsoId=Q99372-4; Sequence=VSP_004246;

CC Name=5; Sequence=VSP_004244, VSP_004245;

CC IsoId=Q99372-5; Sequence=VSP_004244, VSP_004246;

CC Name=6; Sequence=VSP_004245, VSP_004246;

CC IsoId=Q99372-6; Sequence=VSP_004245, VSP_004246;

CC Name=7; Sequence=VSP_004244, VSP_004246;

CC IsoId=Q99372-7; Sequence=VSP_004244, VSP_004246;

CC Name=8; Sequence=VSP_004244, VSP_004245, VSP_004246;

CC IsoId=Q99372-8; Sequence=VSP_004244, VSP_004245, VSP_004246;

CC -1- PTM: The crosslinks are made of deaminated Lys.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

CC EMBL; M0647; AAA2269.1; -

CC EMBL; J04035; AAA2268.1; -

CC EMBL; M86372; AAA42271.1; -

CC EMBL; M86355; AAA42271.1; JOINED.

CC EMBL; M86363; AAA42271.1; JOINED.

CC EMBL; M86364; AAA42271.1; JOINED.

CC EMBL; M86366; AAA42271.1; JOINED.

CC EMBL; M86371; AAA42271.1; JOINED.

CC EMBL; M86376; AAA42272.1; -

CC EMBL; M86373; AAA42272.1; JOINED.

CC EMBL; M86375; AAA42272.1; JOINED.

CC PIR; A36106; EART.

CC InterPro; IPR003979; tropoelastin.

CC PRINTS; PR01500; TROPOELASTIN.

CC Structural protein; Connective tissue; Repeat; Signal;

CC Alternative splicing.

CC NON_TER 1

CC SIGNAL <1 21

CC CHAIN 22 864

CC DISULFID 854 859

CC VARSPPLIC 263 307

CC FT VARSPPLIC 308 308

CC FT VARSPPLIC 308 308

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

CC FT VARSPPLIC 809 823

QY 117 ELMILIAITNLLGQNTPAIVNNAEYGEWAODAAAMEGYAAATATATATLLPPEEAPMT 176
Db 487 -----CAGTLGLGVP-----AVPGALPGAVPCALPGAVPGALP-GAVPGVP 527
QY 177 SAGGL---EQAAAVEEASDTAAANQLMNNVPAALQLOLAQPTQTTTSS-KLGLWKTVS 232
Db 528 GTGGVPGAGTPAAAAAATAA---KAGQYGLGPGVGGVGGVGGVGGVGGVGGV 579
QY 233 PHRSFISNMVSMANNHSMNTNSGVSM-TNTLSSMLKGF-APAAAAQAVQTAAQNGVRAVS 290
Db 580 PG-----GVTGIGTGPCTGLVPGDLGGAGTPAAAKSAKAAKAAKAAQYRAA 624
QY 291 SLGSSSLGSSGLGGV-----AANLGRAASV---GSLSVPOAWAAANQAVTPAARALPLT 341
Db 625 GLGAGVPLGVLGAGVPGFAGAGGFGAGAGVPGFGAGAVPGSLAASAKAAKYGAAGGL--- 681
QY 342 SLTSAARERGPQOMLGGPLVQMGARAG-GGLSGV 374
Db 682 -----GGPGGLGPGGLGGPGGGPGGGPGGLGGV 708

Search completed: August 25, 2004, 03:10:58
Job time : 28 secs

Query Match 7.7%; Score 150; DB 1; Length 864;
Best Local Similarity 25.9%; Pred. No. 0.11;
Matches 102; Conservative 31; Mismatches 153; Indels 108; Gaps 18;
QY 5 GALP-----PEINSARMYAGCSASILVAAQWDSVADLPSAASAFQSVVWGLTVGSM-- 58
Db 399 GGIPGVGGPGGPGVGGPGAVPAAAKAAKAAK--YGARGGVGIPTVGGAGGPG 456
QY 59 --IGSSAGLWVAASPYVAMSVTAQAEELTAQVRVAAAYETAYGLTVPPVIAENRA 116
Db 457 YGVGAGAGL-----GGASQAAA---AAAAKAAKYAG----- 486

This Page Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:13:09 ; Search time 127 Seconds
(without alignments)
967.510 Million cell updates/sec

Title: US-09-724-685-107
Perfect score: 1949
Sequence: 1 MYDFGALPEINARMYAGP.....SCVLEVPFRPYMHPSPAG 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%
Maximum Watch 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pcp.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1949	100.0	391	12	US-09-886-349A-14
2	1949	100.0	391	14	US-10-193-002-102
3	1949	100.0	391	14	US-10-084-843-107
4	1949	100.0	391	14	US-10-098-732A-14
5	1949	100.0	596	9	US-09-287-849-26
6	1949	100.0	596	12	US-09-886-349A-20
7	1949	100.0	596	14	US-10-359-460-26
8	1949	100.0	596	14	US-10-098-732A-20
9	1949	100.0	600	9	US-09-287-849-22
10	1949	100.0	600	14	US-10-359-460-22
11	1949	100.0	723	15	US-10-369-983-2
12	1949	100.0	729	12	US-09-886-349A-18
13	1949	100.0	729	14	US-10-098-732A-18
14	1949	100.0	729	15	US-10-369-983-21
15	1949	100.0	729	15	US-10-369-983-22

16	1949	100.0	813	15	US-10-369-983-15	Sequence 15, Appl
17	1949	100.0	825	15	US-10-369-983-14	Sequence 14, Appl
18	1949	100.0	875	15	US-10-369-983-13	Sequence 13, Appl
19	1949	100.0	930	14	US-10-098-732A-65	Sequence 65, Appl
20	1949	100.0	930	15	US-10-369-983-12	Sequence 12, Appl
21	1949	100.0	1010	15	US-10-369-983-4	Sequence 4, Appl
22	1949	100.0	1016	15	US-10-369-983-18	Sequence 18, Appl
23	1949	100.0	1022	15	US-10-369-983-17	Sequence 17, Appl
24	1949	100.0	1154	15	US-10-369-983-16	Sequence 16, Appl
25	1949	99.7	729	9	US-09-287-849-2	Sequence 2, Appl
26	1949	99.7	729	12	US-09-886-349A-16	Sequence 16, Appl
27	1949	99.7	729	14	US-10-359-460-2	Sequence 2, Appl
28	1949	99.7	729	14	US-10-098-732A-16	Sequence 2, Appl
29	1949	99.7	729	15	US-10-359-459-2	Sequence 8, Appl
30	1934	99.2	391	12	US-09-872-186-8	Sequence 106, App
31	1652.5	84.8	396	14	US-10-193-002-106	Sequence 111, App
32	1652.5	84.8	396	14	US-10-084-843-111	Sequence 62455, A
33	1583	81.2	393	12	US-10-282-122A-64855	Sequence 64892, A
34	1583	81.2	393	12	US-10-282-122A-64892	Sequence 104, App
35	1486.5	76.3	359	14	US-10-193-002-104	Sequence 109, App
36	1486.5	76.3	359	14	US-10-084-843-109	Sequence 8, Appl
37	1187.5	60.9	358	9	US-09-287-849-8	Sequence 8, Appl
38	1187.5	60.9	358	14	US-10-359-460-8	Sequence 12, Appl
39	1187	60.9	263	12	US-09-886-349A-12	Sequence 92, Appl
40	1187	60.9	263	14	US-10-193-002-92	Sequence 91, Appl
41	1187	60.9	263	14	US-10-084-843-91	Sequence 12, Appl
42	1187	60.9	263	14	US-10-098-732A-12	Sequence 126, App
43	766.5	39.3	400	9	US-09-073-009-126	Sequence 126, App
44	766.5	39.3	400	9	US-09-793-306-126	Sequence 62027, A
45	742.5	38.1	405	12	US-10-282-122A-62027	

ALIGNMENTS

RESULT 1
US-09-886-349A-14
; Sequence 14, Application US/09886349A
; Publication No. US20040086523A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886.349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: MTB39 (TbH9FL)
US-09-886-349A-14

Query Match 100.0%; Score 1949; DB 12; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MYDFGALPEINARMYAGPGSASLVAAQAQMDWSVASDLFSAASAFQSVVWGLTVGSWIG 60
DB	1	MYDFGALPEINARMYAGPGSASLVAAQAQMDWSVASDLFSAASAFQSVVWGLTVGSWIG 60
QY	61	SSAGLNVAAASPYVAMSVTAQAEITAAQVAAAYETAYGLTVPPIVIAENRAELMI 120
DB	61	SSAGLNVAAASPYVAMSVTAQAEITAAQVAAAYETAYGLTVPPIVIAENRAELMI 120

0v 1 MVDEGALPPEIN SARMYAGPGSASLVAAOAMNDVSADLFSAAAFQSVWGLTVGSWIG 60

```

Db      9  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68
QY      61  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120
Db      69  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 128
QY      121  LIATNLLGQNTPTAIVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 180
Db      129  LIATNLLGQNTPTAIVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 188
QY      181  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 240
Db      189  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 248
QY      241  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAAAQAQVTAQAQNGVRAMSSSLGSSG 300
Db      249  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAAAQAQVTAQAQNGVRAMSSSLGSSG 308
QY      301  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 360
Db      309  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 368
QY      361  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 391
Db      369  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 399

```

RESULT 6

```

US-09-886-349A-20
; Sequence 20, Application US/09886349A
; Publication No. US20040086523A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Reed, Steven
; APPLICANT: Alderson, Mark
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009070US
; CURRENT APPLICATION NUMBER: US/09/886,349A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/597,796
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 60/265,737
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein TBH9-Ra35 (designated MTB59P)
US-09-886-349A-20

```

```

Query Match      100.0%; Score 1949; DB 12; Length 596;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
Db      9  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68
QY      61  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120
Db      69  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 128
QY      121  LIATNLLGQNTPTAIVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 180
Db      129  LIATNLLGQNTPTAIVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 188
QY      181  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 240

```

```

Db      189  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 248
QY      241  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAAAQAQVTAQAQNGVRAMSSSLGSSG 300
Db      249  MVSVAANHMTNSGVSMNTLTSSMLKGFAPAAAQAQVTAQAQNGVRAMSSSLGSSG 308
QY      301  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 360
Db      309  LGGGVAANLGRAASVGLSVLPQAWAANAQVTPAARALPLTSLTSAARPGQMLGGLPV 368
QY      361  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 391
Db      369  GQMGARAGGSLGVLVPRPPRYVMPHSPAAG 399

```

RESULT 7

```

US-10-359-460-26
; Sequence 26, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
US-10-359-460-26

```

```

Query Match      100.0%; Score 1949; DB 14; Length 596;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
Db      9  MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 68
QY      61  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120
Db      69  SSAGLMVAASPPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 128
QY      121  LIATNLLGQNTPTAIVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 180
Db      129  LIATNLLGQNTPTAIVNEAEYGENWADAAAFGYAAATATATATLPPPEAPEMTSAGG 188
QY      181  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 240
Db      189  LLEQAAAVEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI SN 248

```


QY 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVTAAQNGVRAMSSLGSSLGSSG 300
Db 249 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVTAAQNGVRAMSSLGSSLGSSG 308
QY 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAABRGPGQMLGGLPV 360
Db 309 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAABRGPGQMLGGLPV 368
QY 361 GQMGARAGGGLSGVLVRPPRPVMPHSPAAG 391
Db 369 GQMGARAGGGLSGVLVRPPRPVMPHSPAAG 399

RESULT 8
US-10-098-732A-20
; Sequence 20, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderman, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/099,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 20
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein TSH9-Ra35 (designated MTB59P)
US-10-098-732A-20

Query Match 100.0%; Score 1949; DB 14; Length 596;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60
Db 9 MVDFGALPPEINSARMYAGPGSASLVAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 68
QY 61 SSAGLMVAAAAPYVAMSVTAGQAEELTAQVRVAAAAYETAYGLTVPVPIAENRAELMI 120
Db 69 SSAGLMVAAAAPYVAMSVTAGQAEELTAQVRVAAAAYETAYGLTVPVPIAENRAELMI 128
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPFEAPEMTSAGG 180
Db 129 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPFEAPEMTSAGG 188
QY 181 LLEQAAAVEASDTPAAANQLMNVFQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240
Db 189 LLEQAAAVEASDTPAAANQLMNVFQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 248
QY 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVTAAQNGVRAMSSLGSSLGSSG 300
Db 249 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVTAAQNGVRAMSSLGSSLGSSG 308
QY 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAABRGPGQMLGGLPV 360
Db 309 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAABRGPGQMLGGLPV 368
QY 361 GQMGARAGGGLSGVLVRPPRPVMPHSPAAG 391
Db 369 GQMGARAGGGLSGVLVRPPRPVMPHSPAAG 399

RESULT 9
US-09-287-849-22
; Sequence 22, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 22
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-09-287-849-22

Query Match 100.0%; Score 1949; DB 9; Length 600;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFGALPPEINSARMYAGPGSASLVAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60
Db 9 MVDFGALPPEINSARMYAGPGSASLVAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 68
QY 61 SSAGLMVAAAAPYVAMSVTAGQAEELTAQVRVAAAAYETAYGLTVPVPIAENRAELMI 120
Db 69 SSAGLMVAAAAPYVAMSVTAGQAEELTAQVRVAAAAYETAYGLTVPVPIAENRAELMI 128
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPFEAPEMTSAGG 180
Db 129 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPFEAPEMTSAGG 188
QY 181 LLEQAAAVEASDTPAAANQLMNVFQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240
Db 189 LLEQAAAVEASDTPAAANQLMNVFQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 248
QY 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVTAAQNGVRAMSSLGSSLGSSG 300
Db 249 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVTAAQNGVRAMSSLGSSLGSSG 308
QY 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAABRGPGQMLGGLPV 360
Db 309 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAABRGPGQMLGGLPV 368
QY 361 GQMGARAGGGLSGVLVRPPRPVMPHSPAAG 391
Db 369 GQMGARAGGGLSGVLVRPPRPVMPHSPAAG 399

RESULT 10
US-10-359-460-22
; Sequence 22, Application US/10359460

Publication No. US20030147911A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
FILE REFERENCE: 014058-009020US
CURRENT APPLICATION NUMBER: US/10/359,460
CURRENT FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US/09/287,849
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: US 08/818,112
PRIOR FILING DATE: 1997-03-13
PRIOR APPLICATION NUMBER: US 08/942,578
PRIOR FILING DATE: 1997-10-01
PRIOR APPLICATION NUMBER: US 09/025,197
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 09/056,556
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: US 09/223,040
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 22
LENGTH: 600
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: tri-fusion
US-10-359-460-22

Query Match 100.0%; Score 1949; DB 14; Length 600;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 60
DB 9 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 68
QY 61 SSAGLMVAASPYVAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
DB 69 SSAGLMVAASPYVAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 128
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 180
DB 129 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 188
QY 181 LLEQAAVEEASDTPAAANQNMNVPOALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 240
DB 189 LLEQAAVEEASDTPAAANQNMNVPOALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 248
QY 241 MYSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300
DB 249 MYSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 308
QY 301 LGGVAAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360
DB 309 LGGVAAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 368
QY 361 GQMGARAGGSLGVLRVPPRPVYMPHSPAAG 391
DB 369 GQMGARAGGSLGVLRVPPRPVYMPHSPAAG 399

RESULT 11
US-10-369-983-2
Sequence 2. Application US/10369983
Publication No. US20030235593A1
GENERAL INFORMATION:

APPLICANT: Skeiky, Yasir
APPLICANT: Guderian, Jeff
APPLICANT: Reed, Steven
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
FILE REFERENCE: 014058-008081US
CURRENT APPLICATION NUMBER: US/10/369,983
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 60/357,351
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 723
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: mutated
OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32MutSA)
US-10-369-983-2

Query Match 100.0%; Score 1949; DB 15; Length 723;
Best Local Similarity 100.0%; Pred. No. 2.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 60
DB 333 MVDGALPPEINARMYAGPGSASLVAAQWDSVADLFSASAFQSVVWGLTVGSGWIG 392
QY 61 SSAGLMVAASPYVAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
DB 393 SSAGLMVAASPYVAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 452
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 180
DB 453 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPFEAPEMTSAGG 512
QY 181 LLEQAAVEEASDTPAAANQNMNVPOALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 240
DB 513 LLEQAAVEEASDTPAAANQNMNVPOALQQLAQTQGTTPSSKLGKLVKTVSPHRSPI SN 572
QY 241 MYSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 300
DB 573 MYSMANNHSMNTSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLSGSSG 632
QY 301 LGGVAAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360
DB 633 LGGVAAANLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 692
QY 361 GQMGARAGGSLGVLRVPPRPVYMPHSPAAG 391
DB 693 GQMGARAGGSLGVLRVPPRPVYMPHSPAAG 723

RESULT 12
US-09-886-349A-18
Sequence 18. Application US/09886349A
Publication No. US20040086523A1
GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir
APPLICANT: Reed, Steven
APPLICANT: Alderson, Mark
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
FILE REFERENCE: 014058-009070US
CURRENT APPLICATION NUMBER: US/09/886,349A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 09/597,796
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US 60/265,737
PRIOR FILING DATE: 2001-02-01
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.1

```

; SEQ ID NO 18
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MTB72FmutSA
; OTHER INFORMATION: (Ral2-TbHp-Ra3mutSA)
; US-09-886-349A-18

Query Match      100.0%; Score 1949; DB 12; Length 729;
Best Local Similarity 100.0%; Pred. No. 2.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 MVDFGALPPEINSARMYAGCSASILVAAAQWDSVASDLFSAASAFQSVVWGLITVGSNIG 60
Db 142 MVDFFGALPPEINSARMYAGPSGSILVAAAQWDSVASDLFSAASAFQSVVWGLITVGSNIG 201

Qy 61 SSAGLMWAAAASPYVWAMSVTGAQELTAAQVRVAAAAAYETAYGLTVPPFVIAENRAELMI 120
Db 202 SSAGLMWAAAASPYVWAMSVTGAQELTAAQVRVAAAAAYETAYGLTVPPFVIAENRAELMI 261

Qy 121 LIATNLLGONTFAIAVNEAEYGEWKAODAAAMFCYAAATATATATATLLPFEAPEMTSAGG 180
Db 262 LIATNLLGONTFAIAVNEAEYGEWKAODAAAMFCYAAATATATATLLPFEAPEMTSAGG 321

Qy 181 LLEQAAAVEEASDTHAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240
Db 322 LLEQAAAVEEASDTHAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 381

Qy 241 MYSMANNHMSMTNSGVSMNTLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSILGSSG 300
Db 382 MYSMANNHMSMTNSGVSMNTLSSMLKGFAPAAAQAQVOTAAQNGVRAMSSLGSSILGSSG 441

Qy 301 LGGGVAANI GRAASVGSLSVPOWAAAANQAVTPAARALPLTISLTSAAERGPQMLGGLPV 360
Db 442 LGGGVAANI GRAASVGSLSVPOWAAAANQAVTPAARALPLTISLTSAAERGPQMLGGLPV 501

Qy 361 GQMGAPAGGGLSGVLVRPPRPYVWPHSPAAG 391
Db 502 GQMGAPAGGGLSGVLVRPPRPYVWPHSPAAG 532

```

```

RESULT 13
US-10-098-732A-18
; Sequence 18, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72FMutSA
; OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)
US-10-098-732A-18

```

Query Match	100.0%	Score 1349;	DB 14;	Length 729;
Best Local Similarity	100.0%			
Matches 31:	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

[illegible]

```

RESULT 14
US-10-369-983-21
; Sequence 21, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72F
US-10-369-983-21

```

```
Query Match      100.0%; Score 1949; DB 15; Length 729;
Best Local Similarity 100.0%; Pred. No. 2.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1	MVDFGALPPEINSARMYAGPGSASLVAQAQWDSVASDLFSAASAFQSVVWGLTVGWSMG	60
		
Db	142	MVDFGALPPEINSARMYAGPGSASLVAQAQWDSVASDLFSAASAFQSVVWGLTVGWSMG	201
		
Qy	61	SSAGLMVAAAAPYVWMSVTAGQAELTAAQVRVAAAAYETAYGLTVPVPIAENAEALMI	120
		
Db	202	SSAGLMVAAAAPYVWMSVTAGQAELTAAQVRVAAAAYETAYGLTVPVPIAENAEALMI	261
		
Qy	121	LIATNLLGONTPATAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPENTSAGG	180
		
Db	262	LIATNLLGONTPATAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPENTSAGG	321
		
Qy	181	LLEQAAAYEEASDTAAANQLMNNVPQALQQLAQPTQGTTPSKLGGMLKNTVSPHSPISN	240
		
Db	322	LLEQAAAYEEASDTAAANQLMNNVPQALQQLAQPTQGTTPSKLGGMLKNTVSPHSPISN	381

Search completed: August 25, 2004, 03:28:09
Job time : 131 secs

QY 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSLGSSG 300
Db |
382 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSLGSSG 441
QY 301 LGGGVAANLGRAASVGSLSVQAWAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
Db |
442 LGGGVAANLGRAASVGSLSVQAWAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 501
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391
Db |
502 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 532

RESULT 15
US-10-369-983-22
; Sequence 22, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated
; OTHER INFORMATION: MTB72f-mutSA (Mtb72f-mutSA)
US-10-369-983-22

Query Match 100.0%; Score 1949; DB 15; Length 729;
Best Local Similarity 100.0%; Pred. No. 2.6e-145;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDPGALPPEINSGARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 60
Db |
142 MVDPGALPPEINSGARMYAGPGSASLVAAAQWDSVASDLFSAASAFQSVVWGLTVGSWIG 201
QY 61 SSAGLMVAASPPYVAMSVTAGQAEHTAAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120
Db |
202 SSAGLMVAASPPYVAMSVTAGQAEHTAAQVRVAAAAYETAYGLTVPPPIAENRAELMI 261
QY 121 LIATNLILGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLPEEAPEMTSAGG 180
Db |
262 LIATNLILGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLPEEAPEMTSAGG 321
QY 181 LLECAAAVEEASPTAAANQLMNNVPQALQQAQTOQTTPSSKLGGLWKTSPHRSPISN 240
Db |
322 LLECAAAVEEASPTAAANQLMNNVPQALQQAQTOQTTPSSKLGGLWKTSPHRSPISN 381
QY 241 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSLGSSG 300
Db |
382 MYSMANNHMTNSGVSMNTLSSMLKGFAPAAAQAVQTAQAQNGVRAMSSLGSSLGSSG 441
QY 301 LGGGVAANLGRAASVGSLSVQAWAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
Db |
442 LGGGVAANLGRAASVGSLSVQAWAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 501
QY 361 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391
Db |
502 GQMGARAGGGLSGVLRVPPRPYVMPHSPAAG 532

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 02:57:18 ; Search time 128 Seconds
(without alignments)
863.094 Million cell updates/sec

Title: US-09-724-685-107
Perfect score: 1949
Sequence: 1 MVDPCALPPEINSARMYAGP.....SGVLVPPRPVYMPHSPAAG 391.

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1949	100.0	391	2 AAW32381	Aw32381 Mycobacte
2	1949	100.0	391	2 AAW32449	Aw32449 Mycobacte
3	1949	100.0	391	2 AAW43335	Aw43335 Mycobacte
4	1949	100.0	391	2 AAW51702	Aw51702 M. tuberc
5	1949	100.0	391	2 AAY04778	Ay04778 Mycobacte
6	1949	100.0	391	2 AAY38989	Ay38989 M. tuberc
7	1949	100.0	391	2 AAY39132	Ay39132 M. tuberc
8	1949	100.0	391	4 AAU01888	Au01888 M. tuberc
9	1949	100.0	391	5 AAE29707	Ae29707 Mycobacte
10	1949	100.0	391	5 AAE17571	Ae17571 Mycobacte
11	1949	100.0	596	2 AAY32070	Ay32070 Mycobacte
12	1949	100.0	596	5 AAE29710	Ae29710 Mycobacte
13	1949	100.0	596	5 AAE17574	Ae17574 Mycobacte
14	1949	100.0	599	5 AAU74599	Au74599 Antigenic
15	1949	100.0	600	2 AAY32068	Ay32068 Mycobacte
16	1949	100.0	600	5 AAU74597	Au74597 Antigenic
17	1949	100.0	723	7 ADA26354	Ad26354 Mycobacte
18	1949	100.0	729	4 AAO22142	Ao22142 Ra12-H9-3
19	1949	100.0	729	5 AAE29709	Ae29709 Mycobacte
20	1949	100.0	729	5 AAE17573	Ae17573 Mycobacte
21	1949	100.0	729	7 ADA26374	Ad26374 Mycobacte
22	1949	100.0	744	4 AAU01902	Au01902 M. tuberc
23	1949	100.0	813	7 ADA26367	Ad26367 Mycobacte
24	1949	100.0	815	4 AAU01904	Au01904 M. tuberc
25	1949	100.0	825	7 ADA26366	Ad26366 Mycobacte

26	1949	100.0	875	7 ADA26365	Ad26365 Mycobacte
27	1949	100.0	930	5 AAE29731	Ae29731 Mycobacte
28	1949	100.0	930	7 ADA26364	Ad26364 Mycobacte
29	1949	100.0	1010	7 ADA26356	Ad26356 Mycobacte
30	1949	100.0	1016	7 ADA26370	Ad26370 M. bovis
31	1949	100.0	1022	7 ADA26369	Ad26369 Mycobacte
32	1949	100.0	1154	7 ADA26368	Ad26368 Mycobacte
33	1945	99.8	788	4 AAU01903	Au01903 M. tuberc
34	1944	99.7	394	2 AAY04779	Ay04779 Mycobacte
35	1944	99.7	729	5 AAE29708	Ae29708 Mycobacte
36	1944	99.7	729	5 AAE17572	Ae17572 Mycobacte
37	1939	99.5	723	7 ADA26373	Ad26373 Mycobacte
38	1931	99.1	729	2 AAY32059	Ay32059 Mycobacte
39	1902.5	97.6	726	5 AAU74588	Au74588 Antigenic
40	1652.5	84.8	396	2 AAW64337	Aw64337 Mycobacte
41	1652.5	84.8	396	2 AAW81704	Aw81704 M. tuberc
42	1652.5	84.8	396	2 AAY38991	Ay38991 M. tuberc
43	1652.5	84.8	396	2 AAY39134	Ay39134 M. tuberc
44	1583	81.2	393	6 ABU36968	Abu36968 Protein e
45	1583	81.2	393	6 ABU34531	Abu34531 Protein e

ALIGNMENTS

RESULT 1
AAW32381
ID AAW32381 standard; protein; 391 AA.
XX
AC AAW32381;
XX
DT 13-JAN-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen TbH-9FL.
XX
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9709429-A2.
XX
PD 13-MAR-1997.
XX
PF 30-AUG-1996; 96WO-US014675.
XX
PR 01-SEP-1995; 95US-00523435.
PR 22-SEP-1995; 95US-00532136.
PR 22-MAR-1996; 96US-00620280.
PR 05-JUN-1996; 96US-00658800.
PR 12-JUL-1996; 96US-00680573.
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX
XX WPI: 1997-192904/17.
DR N-PSDB; AAT91455.
XX
XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
XX useful for diagnosis of M. tuberculosis infection.
XX
XX Example 3; Page 150-152; 190pp; English.

XX A new immunogenic polypeptide has been developed comprising an
XX immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
XX variant differing only in conservative substitutions and/or
XX modifications). The present sequence represents a M.tuberculosis antigen,
XX TbH-9FL The immunogenic polypeptide can be used to diagnose
XX M.tuberculosis infection by forming complexes with specific antibodies in
XX the sample. Fragments of DNA encoding the immunogenic polypeptide can be
XX used as diagnostic primers or probes and agents that bind to the antigen,
XX

CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
 CC also used for diagnosis

XX Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFOSVVMGLTVGSMIG 60
 DB 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFOSVVMGLTVGSMIG 60

QY 61 SSAGLWVAASPYVAMSVTAGQAEELTAQVRVAAAYETAYGLTVPPVIAENRAELMI 120
 DB 61 SSAGLWVAASPYVAMSVTAGQAEELTAQVRVAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLTPFEAPEMTSAGG 180
 DB 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLTPFEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPAQLQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240
 DB 181 LLEQAAVEEASDTAAANQLMNNVPAQLQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240

QY 241 MVSMAHHMNTNGSVMTNTLSMLKGFAPAAAAQAVQTAQNGVRAMSSLGSSLGSSG 300
 DB 241 MVSMAHHMNTNGSVMTNTLSMLKGFAPAAAAQAVQTAQNGVRAMSSLGSSLGSSG 300

QY 301 LGGVVAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360
 DB 301 LGGVVAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391
 DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

RESULT 2

AAW32449
 ID AAW32449 standard; protein; 391 AA.

AC AAW32449;

XX 09-JAN-1998 (first entry)

DE Mycobacterium tuberculosis antigen TbH-9FL.

KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

XX skin testing; M.tuberculosis.

OS Mycobacterium tuberculosis.

XX WO9709428-A2.

XX 13-MAR-1997.

PF 30-AUG-1996; 96WO-US014674.

XX 01-SEP-1995; 95US-00523436.

PR 22-SEP-1995; 95US-00533634.

PR 22-MAR-1996; 96US-00620874.

PR 05-JUN-1996; 96US-00659683.

PR 12-JUL-1996; 96US-00680574.

XX (CORI-) CORIXA CORP.

PA Reed SG, Skeiky VA, Dillon DC, Campos-Neto A, Houghton R;

XX Vedvick TH, Twardzik DR;

XX WPI; 1997-192903/17.

DR N-PSDB; AAT91521.

PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
 PT useful in vaccines for prevention or treatment of tuberculosis, also for
 XX diagnosis.

PS Example 3; Page 138-139; 168pp; English.

XX A new immunogenic polypeptide has been developed comprising an
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
 CC variant differing only in conservative substitutions and/or
 CC modifications). The present sequence represents a M.tuberculosis antigen,
 CC TbH-9FL. The immunogenic protein, and fusion proteins containing one or
 CC more of the proteins or one of the proteins plus ESAM-6, are useful in
 CC vaccines, preferably when formulated with a non-specific adjuvant, to
 CC induce an immune response against M.tuberculosis (for treatment or
 CC prevention)

XX Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;

Best Local Similarity 100.0%; Pred. No. 1.2e-142;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFOSVVMGLTVGSMIG 60

DB 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFOSVVMGLTVGSMIG 60

QY 61 SSAGLWVAASPYVAMSVTAGQAEELTAQVRVAAAYETAYGLTVPPVIAENRAELMI 120

DB 61 SSAGLWVAASPYVAMSVTAGQAEELTAQVRVAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATLTPFEAPEMTSAGG 180

DB 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATLTPFEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPAQLQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240

DB 181 LLEQAAVEEASDTAAANQLMNNVPAQLQAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240

QY 241 MVSMAHHMNTNGSVMTNTLSMLKGFAPAAAAQAVQTAQNGVRAMSSLGSSLGSSG 300

DB 241 MVSMAHHMNTNGSVMTNTLSMLKGFAPAAAAQAVQTAQNGVRAMSSLGSSLGSSG 300

QY 301 LGGVVAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360

DB 301 LGGVVAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERGPQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

DB 361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391

RESULT 3

AAW64335

ID AAW64335 standard; protein; 391 AA.

XX AAW64335;

AC AAW64335;

XX 17-OCT-2003 (revised)

DT 09-NOV-1998 (first entry)

XX Mycobacterium tuberculosis antigen TbH-9FL.

OS Tuberculosis; infection; diagnosis; antigen; TbH-9FL.

XX Mycobacterium tuberculosis; strain H37Rv.

XX WO9816645-A2.

PN 23-APR-1998.

XX 07-OCT-1997; 97WO-US018214.

XX 11-OCT-1996; 96US-00729622.

PR	13-MAR-1997;	97US-00818111.	
XX	(CORI-) CORIXA CORP.		
XX	Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;		
XX	Vedvick TS, Twardzik DR, Lodes MJ;		
XX	WPI; 1998-251292/22.		
DR	N-PSDB; AAV44395.		
XX			
XX	New isolated Mycobacterium tuberculosis polypeptides and DNA - used to		
PT	develop products for the detection of M. tuberculosis infection and		
PT	diagnosis of tuberculosis.		
XX			
PS	Example 3; Page 133-135; 250pp; English.		
XX			
CC	This polypeptide comprises Mycobacterium tuberculosis antigen TBH-9FL. It		
CC	is encoded by genomic DNA (see AAV44395) isolated from a M. tuberculosis		
CC	strain H37Rv genomic library using a probe from clone TBH-9 (see		
CC	AAV44371). The invention relates to compositions and methods for		
CC	diagnosing tuberculosis. It provides polypeptides (see AAV64291-W64379)		
CC	comprising an antigenic portion of a soluble M. tuberculosis antigen, or		
CC	an immunogenic portion of an M. tuberculosis antigen, as well as DNA		
CC	sequences encoding such polypeptides, recombinant expression vectors and		
CC	transformed or transfected host cells. Also claimed are methods and		
CC	diagnostic kits for detecting M. tuberculosis infection in a patient		
CC	using these polypeptides, antibodies or oligonucleotide probes and		
CC	primers, for the diagnosis of tuberculosis. (Updated on 17-OCT-2003 to		
CC	standardise OS field)		
XX			
SQ	Sequence 391 AA;		
	Query Match 100.0%; Score 1949; DB 2; Length 391;		
	Best Local Similarity 100.0%; Pred. No. 1.2e-142;		
	Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MVDFGALPPPEINSAARMYAGPGSASLVAAQWDSVASDLFSAASATQSVVWGLTVGSGWIG 60		
DB	1 MVDFGALPPPEINSAARMYAGPGSASLVAAQWDSVASDLFSAASATQSVVWGLTVGSGWIG 60		
QY	61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120		
DB	61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120		
QY	121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPPEEAPMTSAGG 180		
DB	121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPPEEAPMTSAGG 180		
QY	181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240		
DB	181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240		
QY	241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSLGSSG 300		
DB	241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSLGSSG 300		
QY	301 LGGGVAANLGRAASVGLSVPOQAAVAAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360		
DB	301 LGGGVAANLGRAASVGLSVPOQAAVAAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360		
QY	361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391		
DB	361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391		
RESULT 4			
AAW81702			
ID	AAW81702 standard; protein; 391 AA.		
XX			
AC	AAW81702;		
XX			
DT	27-JAN-1999 (first entry)		
XX			

DE	M. tuberculosis immunogenic polypeptide TBH-9FL.		
XX	Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;		
XX	vaccine; pharmaceutical; infection; diagnosis.		
XX			
OS	Mycobacterium tuberculosis.		
XX			
PN	WO9816646-A2.		
XX			
PD	23-APR-1998.		
XX			
PF	07-OCT-1997; 97WO-US018293.		
XX			
PR	11-OCT-1996; 96US-00730510.		
XX			
PR	13-MAR-1997; 97US-00818112.		
XX			
XX	(CORI-) CORIXA CORP.		
PA	Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;		
XX	Vedvick TS, Twardzik DR, Lodes MJ;		
PI	WPI; 1998-261042/23.		
XX			
DR	N-PSDB; AAV64503.		
XX			
XX	Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to		
PT	develop products for the detection of M. tuberculosis infection and for		
PT	diagnosis, treatment and prevention of tuberculosis.		
XX			
PS	Example 3B; Page 128-129; 230pp; English.		
XX			
CC	This sequence represents an immunogenic portion of a soluble		
CC	Mycobacterium tuberculosis (MT) antigen which can be used in a method for		
CC	inducing protective immunity against tuberculosis (TB). This sequence can		
CC	be formulated into vaccines and/or pharmaceutical compositions for		
CC	immunising against M. tuberculosis infection or may be used for the		
CC	diagnosis of tuberculosis		
XX			
SQ	Sequence 391 AA;		
	Query Match 100.0%; Score 1949; DB 2; Length 391;		
	Best Local Similarity 100.0%; Pred. No. 1.2e-142;		
	Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MVDFGALPPPEINSAARMYAGPGSASLVAAQWDSVASDLFSAASATQSVVWGLTVGSGWIG 60		
DB	1 MVDFGALPPPEINSAARMYAGPGSASLVAAQWDSVASDLFSAASATQSVVWGLTVGSGWIG 60		
QY	61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120		
DB	61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIVIAENRAELMI 120		
QY	121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180		
DB	121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180		
QY	181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240		
DB	181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQPTQGTTPSSKLGGLWKTVPSPHSPISN 240		
QY	241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSLGSSG 300		
DB	241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSLGSSG 300		
QY	301 LGGGVAANLGRAASVGLSVPOQAAVAAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360		
DB	301 LGGGVAANLGRAASVGLSVPOQAAVAAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360		
QY	361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391		
DB	361 GQMGARAGGGLSGVLRVPPRPVMPHSPAAG 391		
RESULT 5			

AA04778
 ID AAY04778 standard; protein; 391 AA.
 AC AAY04778;
 XX
 DT 06-JUL-1999 (first entry)
 XX
 DE Mycobacterium species protein sequence 5R.
 XX
 KW Secreted protein; Mycobacterium; primer: PCR; amplification; probe;
 KW hybridisation; detection; vaccine; immunisation; infection.
 XX
 OS Mycobacterium sp.
 XX
 PN WO9909186-A2.
 XX
 PD 25-FEB-1999.
 XX
 PF 14-AUG-1998; 98WO-FR001813.
 XX
 PR 14-AUG-1997; 97FR-00010404.
 PR 11-SEP-1997; 97FR-00011325.
 XX
 PA (INSP) INST PASTEUR.
 XX
 XX Gicquel B, Portnoie D, Lim E, Pelicic V, Guigueno A;
 PI Goguet de La Salmoniere Y;
 XX
 DR WPI; 1999-181045/15.
 DR N-PSDB; AAX34030.
 XX
 PT Mycobacterial DNA vectors containing reporter constructs - for
 PT identifying coding or promoter sequences involved in infection-associated
 PT protein expression.
 XX
 PS Claim 32; Fig 5R; 309pp; French.
 XX
 CC Sequences AAY04742-Y05000 and AAY07201-Y07204 represent secreted proteins
 CC from various Mycobacterium species microorganisms. The encoding
 CC nucleotide sequences can be used as primers and probes for methods for
 CC detecting and identifying mycobacteria, especially belonging to the M.
 CC tuberculosis complex. The encoded proteins can be used in vaccines for
 CC immunisation against a bacterial or viral infection
 XX
 SQ Sequence 391 AA;
 Query Match 100.0%; Score 1949; DB 2; Length 391;
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDGALPPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60
 DB 1 MVDGALPPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60
 QY 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
 DB 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
 QY 121 LIATNLGQNTPAIAVNEAEYGEWQAADAAAMFYAAATATATATATLLPFEAPMTSAGG 180
 DB 121 LIATNLGQNTPAIAVNEAEYGEWQAADAAAMFYAAATATATATATLLPFEAPMTSAGG 180
 QY 181 LLEQAAAVEASDTAAANQLMNNVPQALQQAQTOGTTTSSKLGGLWKTVPSPHSPISN 240
 DB 181 LLEQAAAVEASDTAAANQLMNNVPQALQQAQTOGTTTSSKLGGLWKTVPSPHSPISN 240
 QY 241 MVSNNHNSMTNSGVSMTNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSLGSSG 300
 DB 241 MVSNNHNSMTNSGVSMTNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSLGSSG 300
 QY 301 LGGGVAANTLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAABRGPCQMLGGLPV 360
 DB 301 LGGGVAANTLGRAASVGSLSVPQAAANQAVTPAARALPLTSLTSAABRGPCQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPRPYVMPHSPAAG 391
 DB 361 GQMGARAGGGLSGVLRVPRPYVMPHSPAAG 391

RESULT 6

AAV38989
 ID AAY38989 standard; protein; 391 AA.
 XX
 AC AAY38989;
 XX
 DT 05-NOV-1999 (first entry)
 XX
 DE M. tuberculosis recombinant antigen protein TBH-9FL.
 XX
 KW Antigen; diagnosis; detection; infection; antibody; immunisation;
 KW vaccine; immunity.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO9942118-A2.
 XX
 PD 26-AUG-1999.
 XX
 PF 17-FEB-1999; 99WO-US003265.
 PR 18-FEB-1998; 98US-00024753.
 PR 05-MAY-1998; 98US-00072596.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
 XX
 DR WPI; 1999-527416/44.
 DR N-PSDB; AAZ19093.
 XX
 PT New polypeptide comprising antigenic portions of M. tuberculosis.
 PS Example 3; Page 168-169; 323pp; English.
 CC This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against M.
 CC tuberculosis infection. The new detection methods are needed as current
 CC vaccination strategies do not provide 100% immunity
 XX
 SQ Sequence 391 AA;

Query Match 100.0%; Score 1949; DB 2; Length 391;
 Best Local Similarity 100.0%; Pred. No. 1.2e-142;
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDGALPPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60
 DB 1 MVDGALPPPEINARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSGWG 60
 QY 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
 DB 61 SSAGLMVAASPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPIAENRAELMI 120
 QY 121 LIATNLGQNTPAIAVNEAEYGEWQAADAAAMFYAAATATATATLLPFEAPMTSAGG 180
 DB 121 LIATNLGQNTPAIAVNEAEYGEWQAADAAAMFYAAATATATATLLPFEAPMTSAGG 180
 QY 181 LLEQAAAVEASDTAAANQLMNNVPQALQQAQTOGTTTSSKLGGLWKTVPSPHSPISN 240
 DB 181 LLEQAAAVEASDTAAANQLMNNVPQALQQAQTOGTTTSSKLGGLWKTVPSPHSPISN 240
 QY 241 MVSNNHNSMTNSGVSMTNTLSSMLKGFAPAAAQAVQTAQNGVRAMSSLGSSLGSSG 300


```

Db      241  MYSMANNHSMNTNSGVSMNTLSSMLKGFAPAAAAQAVOTTAQNGVRAMSSLGSSGSSG 300
QY      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPQOMLGGJLV 360
Db      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPQOMLGGJLV 360
QY      361  GQMGARAGGGLSGVLRVPPRPYVPHSPAAG 391
Db      361  GQMGARAGGGLSGVLRVPPRPYVPHSPAAG 391

RESULT 7
AAU01888
ID      AAY39132 standard; protein; 391 AA.
XX
AC      AAY39132;
XX
DT      05-NOV-1999 (first entry)
XX
DE      M. tuberculosis antigen Tbh-9FL amino acid sequence.
XX
KW      Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW      immunotherapy; diagnosis; immunisation; vaccine; infection;
KW      immune response; skin test.
XX
OS      Mycobacterium tuberculosis.
XX
FN      WO9942076-A2.
XX
PD      26-AUG-1999.
XX
PF      17-FEB-1999; 99WO-US003268.
XX
PR      18-FEB-1998; 98US-00025197.
PR      05-MAY-1998; 98US-00072967.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI      Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
DR      WPI; 1999-527409/44.
DR      N-PSDB; AAZ19305.
XX
PT      New antigens from Mycobacterium tuberculosis useful in diagnostic skin
PT      tests and protective or therapeutic vaccines or compositions.
PS      Example 3; Page 123-124; 299pp; English.
XX
XX      The present invention describes polypeptides comprising an immunogenic
XX      part of a Mycobacterium tuberculosis antigen (Ag). Also described are
XX      vaccines and fusion protein containing M. tuberculosis Ag's. M.
XX      tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
XX      polypeptides fragments, can be used in pharmaceutical compositions or
XX      vaccines to generate a protective or therapeutic immune response to M.
XX      tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
XX      Ag can induce proliferation of, or cytokine secretion by, T, B or natural
XX      killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
XX      to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of
XX      the present invention.
XX
SQ      Sequence 391 AA;
Query Match      100.0%; Score 1949; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.2e-142;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1  MVDFGALPPEINARMYAGPQSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSWIG 60
Db      1  MVDFGALPPEINARMYAGPQSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSWIG 60
QY      61  SSAGLWVAASPYYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPIAENRAELMI 120

```

```

Db      61  SSAGLWVAASPYYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPIAENRAELMI 120
QY      121  LIATNLGQNTPAIYNAEYGEWMAQDRAAFGYAAATATATATATALLPPEAEPMTSAGG 180
Db      121  LIATNLGQNTPAIYNAEYGEWMAQDRAAFGYAAATATATATATALLPPEAEPMTSAGG 180
QY      181  LLEQAAAAVEEASDTAAANQLMNNVPOALQLOAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240
Db      181  LLEQAAAAVEEASDTAAANQLMNNVPOALQLOAQPTQGTTPSSKLGGLWKTVPSPHRSPI 240
QY      241  MVSMAHHMNTNSGVSMNTLSSMLKGFAPAAAAQAVOTTAQNGVRAMSSLGSSGSSG 300
Db      241  MVSMAHHMNTNSGVSMNTLSSMLKGFAPAAAAQAVOTTAQNGVRAMSSLGSSGSSG 300
QY      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPQOMLGGJLV 360
Db      301  LGGGVAANLGRAASVGSLSVPOQAAAAQAVTPAARALPLTSLTSAARGPQOMLGGJLV 360
QY      361  GQMGARAGGGLSGVLRVPPRPYVPHSPAAG 391
Db      361  GQMGARAGGGLSGVLRVPPRPYVPHSPAAG 391

RESULT 8
AAU01888
ID      AAU01888 standard; protein; 391 AA.
XX
AC      AAU01888;
XX
DT      29-AUG-2001 (first entry)
XX
DE      M. tuberculosis antigen Tbh9 (Mtb39A).
XX
KW      Tbh9; Mtb39A; antigen; vaccine; tuberculosis; AIDS;
KW      acquired immunodeficiency disease.
XX
OS      Mycobacterium tuberculosis.
XX
FN      WO200124820-A1.
XX
PD      12-APR-2001.
XX
PF      10-OCT-2000; 2000WO-US028095.
XX
PR      07-OCT-1999; 99US-0158338P.
PR      07-OCT-1999; 99US-0158425P.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Skeiky Y, Reed S, Houghton RL, Mcneill PD, Dillon DC, Lodes ML;
XX
XX      WPI; 2001-290576/30.
XX      N-PSDB; AAS03779.
XX
PT      Vaccinating against Mycobacteria infections in mammals using fusion
PT      proteins comprising combinations of heterologous antigens.
XX
PS      Example 2; Page 151-152; 169pp; English.
XX
XX      The sequence represents Mycobacterium tuberculosis Tbh9 (also known as
XX      Mtb39A), an M. tuberculosis antigen. Compositions comprising at least 2
XX      heterologous antigens, as a fusion protein, and vectors expressing the
XX      fusion proteins are used as vaccines to prophylactically immunise mammals
XX      (especially humans) against infection by Mycobacteria. The compositions
XX      contain at least 2 heterologous antigens that increase the serological
XX      sensitivity of individuals infected with tuberculosis, a disease
XX      frequently affecting patients with acquired immunodeficiency disease,
XX      AIDS
XX
SQ      Sequence 391 AA;
Query Match      100.0%; Score 1949; DB 4; Length 391;

```

Best Local Similarity 100.0%; Pred. No. 1.2e-142; Mismatches 0; Indels 0; Gaps 0;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFGALPPEINSAARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60
DB 1 MVDFGALPPEINSAARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60

QY 61 SSAGLMVAASAPYVAMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLMVAASAPYVAMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLGONTPTAIVNEAEGEMWAQDAAMFGYAAATATATATLTPPEEAPEMTSAGG 180
DB 121 LIATNLGONTPTAIVNEAEGEMWAQDAAMFGYAAATATATATLTPPEEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQNMNVPOALQOQAQPTQGTTPSSKLGGLWKTVPSPHRSPISN 240
DB 181 LLEQAAVEEASDTAAANQNMNVPOALQOQAQPTQGTTPSSKLGGLWKTVPSPHRSPISN 240

QY 241 MVSMAHNMSTNGSVMTNTLSMLKGFAPAAAQVQTAQNGVRAMSSLSGSSG 300
DB 241 MVSMAHNMSTNGSVMTNTLSMLKGFAPAAAQVQTAQNGVRAMSSLSGSSG 300

QY 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERCPGQMLGGLPV 360
DB 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERCPGQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391
DB 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391

RESULT 9
AAE29707
ID AAE29707 standard; protein; 391 AA.
XX AAE29707;
XX
XX
XX 27-JAN-2003 (first entry)
XX
XX Mycobacterium sp. TbH9FL antigenic protein.
XX Vaccine; immunity; diagnostic agent; gene therapy; TbH9FL antigen.
XX Mycobacterium sp.
XX
XX WO200272792-A2.
XX
XX 19-SEP-2002.
XX
XX 13-MAR-2002; 2002WO-US008223.
XX
XX 13-MAR-2001; 2001US-0275837P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Brannon M, Guderian J;
XX
XX WPI; 2002-759844/82.
XX N-PSDB; AAD47082.
XX
XX New recombinant nucleic acid molecule comprising a Leishmania TSA, Leif, M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity against pathogenic microorganisms e.g. Leishmania and Mycobacterium tuberculosis.
XX
XX Disclosure; Page 86-87; 155pp; English.
XX
XX The invention relates to a recombinant nucleic acid molecule encoding a fusion polypeptide. The recombinant nucleic acid comprises a heterologous polynucleotide sequence encoding an antigen or an antigenic fragment from Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a polypeptide or its fragment. The Leishmania polynucleotide is selected

CC from TSA, Leif, M15, and 6H polynucleotides. Sequences of the invention are used in methods for eliciting immune response in mammals. They are useful as vaccines to elicit protective immunity against pathogenic microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion polypeptides are used for enhancing the expression of polynucleotides, as in vivo diagnostic agents and for raising antibodies in a non-human animal. The invention is used in gene therapy. The present sequence is CC Mycobacterium sp. TbH9FL antigenic protein
XX
XX Sequence 391 AA;
SQ

Query Match 100.0%; Score 1949; DB 5; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.2e-142;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFGALPPEINSAARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60
DB 1 MVDFGALPPEINSAARMYAGPGSASLVAAAQWDSVADLFSAAAFOSVWGLTVGSGWIG 60

QY 61 SSAGLMVAASAPYVAMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLMVAASAPYVAMSVTAGQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

QY 121 LIATNLGONTPTAIVNEAEGEMWAQDAAMFGYAAATATATLTPPEEAPEMTSAGG 180
DB 121 LIATNLGONTPTAIVNEAEGEMWAQDAAMFGYAAATATATLTPPEEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQNMNVPOALQOQAQPTQGTTPSSKLGGLWKTVPSPHRSPISN 240
DB 181 LLEQAAVEEASDTAAANQNMNVPOALQOQAQPTQGTTPSSKLGGLWKTVPSPHRSPISN 240

QY 241 MVSMAHNMSTNGSVMTNTLSMLKGFAPAAAQVQTAQNGVRAMSSLSGSSG 300
DB 241 MVSMAHNMSTNGSVMTNTLSMLKGFAPAAAQVQTAQNGVRAMSSLSGSSG 300

QY 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERCPGQMLGGLPV 360
DB 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERCPGQMLGGLPV 360

QY 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391
DB 361 GQMGARAGGGLSGVLRVPPRPYPVPHSPAAG 391

RESULT 10
AAE17571
ID AAE17571 standard; protein; 391 AA.
XX AAE17571;
XX
XX 22-APR-2002 (first entry)
XX
XX Mycobacterium species MTB39 (TbH9) protein #2.
XX
XX Fusion protein; antigen; serological sensitivity; immune response; tuberculosis; infection; vaccine; MTB39; TbH9 protein.
XX Mycobacterium sp.
XX WO200198460-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US019959.
XX
XX 20-JUN-2000; 2000US-00597796.
XX 01-FEB-2001; 2001US-0265737P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky Y, Reed S, Alderson M;
XX WPI; 2002-147798/19.
XX

DR N-PSDB; AAD28341.
XX Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 83; Page 102-103; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is Mycobacterium species
CC MTB39 (TbH9) protein
XX
SQ Sequence 391 AA;
Query Match 100.0%; Score 1949; DB 5; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.2e-142;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSMIG 60
Db 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSMIG 60
QY 61 SSAGLMVAAAAPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
Db 61 SSAGLMVAAAAPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180
Db 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTGTTTSSKLGGLWKTVPSPHSPISN 240
Db 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTGTTTSSKLGGLWKTVPSPHSPISN 240
QY 241 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 300
Db 241 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 300
QY 301 LGGVVAANLGRAASVGLSVPOQAAWAAQAVTPAARALPLTSLTSAERPGQWLGGLPV 360
Db 301 LGGVVAANLGRAASVGLSVPOQAAWAAQAVTPAARALPLTSLTSAERPGQWLGGLPV 360
QY 361 GQMGARAGGGLSVGLRVPPRPYVMPHSPAAG 391
Db 361 GQMGARAGGGLSVGLRVPPRPYVMPHSPAAG 391
RESULT 11
AAY32070
ID AAY32070 standard; protein; 596 AA.
XX
XX AAY32070;
XX
XX 17-JAN-2000 (first entry)
DT
XX Mycobacterium tuberculosis antigen fusion protein Mtb59f.
DE

XX Tuberculosis; antigen; fusion protein; Mtb59f; TbH9; Ra35; diagnosis;
XX therapy; vaccine; immunogen.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT Peptide 1..8
FT Protein /note= "Met/His tag"
FT Protein 9..140
FT Protein /note= "Ra12"
FT Protein 143..596
FT Protein /note= "TbH9"
PN WO9951748-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Alderson M, Campos-Neto A;
XX
XX WPI; 1999-601610/51.
XX N-PSDB; AAZ20205.
XX
XX New fusion proteins useful for diagnosis, prevention and treatment of
XX tuberculosis.
PT
XX Claim 1; Fig 12A-B; 83pp; English.
XX
XX This sequence represents a recombinant Mycobacterium tuberculosis bi-
CC antigen fusion protein, termed Mtb59f, composed of the antigens TbH9 and
CC Ra35. The fusion protein is expressed in host cells using a vector
CC carrying a polynucleotide (see AAZ20205) comprising the coding sequences
CC for the 2 antigens. The invention provides fusion proteins (see AAY32059-
CC 71) containing at least 2 M. tuberculosis antigens. The new fusion
CC proteins and polynucleotides encoding them are useful as vaccines for
CC preventing tuberculosis (claimed), for diagnosis (via in vitro assays or
CC intradermal skin tests for detection of anti-M. tuberculosis antibodies),
CC monitoring of disease progression, and treatment of tuberculosis. They
CC are more effective immunogens than mixtures of the individual protein
CC components
XX
SQ Sequence 596 AA;
Query Match 100.0%; Score 1949; DB 2; Length 596;
Best Local Similarity 100.0%; Pred. No. 2.1e-142;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSMIG 60
Db 9 MVDGALPPEINSARMYAGPGSASLVAAQWDSVADLFSAAAFQSVVWGLTVGSMIG 69
QY 61 SSAGLMVAAAAPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 120
Db 69 SSAGLMVAAAAPYVAMSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVPIAENRAELMI 128
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 180
Db 129 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATALLPPEEAPMTSAGG 188
QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTGTTTSSKLGGLWKTVPSPHSPISN 240
Db 189 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTGTTTSSKLGGLWKTVPSPHSPISN 248
QY 241 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 300
Db 249 MYSMANNHMTNSGVSWTNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLGSSG 308

CC immunised with BCG. They are useful in the diagnosis, treatment and
 CC prevention of Mycobacterium infection. The fusion proteins and the
 CC polynucleotides are useful as diagnostic tools in patients infected with
 CC Mycobacterium, in vitro and in vivo assays for detecting humoral
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the
 CC diagnosis of an infection or monitoring of disease progression, as
 CC immunogens to generate or elicit a protective immune response in a
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human
 CC animal. Sequences of the invention are also used as vaccines. MTB32A
 CC fusion proteins of the invention are useful as in vivo diagnostic agents
 CC for intradermal skin test. The present sequence is Mycobacterium species
 CC MTB59f (TbH9-Ra35) fusion protein
 XX
 XX Sequence 596 AA;

Query Match 100.0%; Score 1949; DB 5; Length 596;
 Best Local Similarity 100.0%; Pred. No. 2.1e-142;
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60
 Db 9 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 68
 QY 61 SSAGLMVAASPYVAVMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120
 Db 69 SSAGLMVAASPYVAVMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 128
 QY 121 LIATNLLGQNTPAIAVNEAEGYEMWAOADAAAFGYAAATATATATLLPFEAPMTSAGG 180
 Db 129 LIATNLLGQNTPAIAVNEAEGYEMWAOADAAAFGYAAATATATATLLPFEAPMTSAGG 188
 QY 181 LLEQAAVEEASDTAAANQLMNVPOALQQAQPTGTTTSSKLGGLWKTVPSPHSPISN 240
 Db 189 LLEQAAVEEASDTAAANQLMNVPOALQQAQPTGTTTSSKLGGLWKTVPSPHSPISN 248
 QY 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 300
 Db 249 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 308
 QY 361 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 391
 Db 369 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 399

RESULT 14
 AAU74599

XX AAU74599 standard; protein; 599 AA.

XX AAU74599;

DT 29-AUG-2003 (revised)

DT 08-MAY-2002 (first entry)

DE Antigenic fusion protein Tb59-Ra35 (Mtb59f).

XX Fusion protein; tuberculosis; Mycobacterium tuberculosis;
 XX tuberculostatic; immunogen; vaccine; Tb59-Ra35; Mtb59f.

XX Mycobacterium tuberculosis.

OS Chimeric.

XX Key Location/Qualifiers

XX Misc-difference 597

FT /label= OTHER

FT /note= "OTHER= Xaa. Xaa= In frame stop codon"

XX US2002009459-A1.

XX 24-JAN-2002.

XX 07-APR-1999; 99US-00287849.
 XX 13-MAR-1997; 97US-00818112.
 PR 01-OCT-1997; 97US-00942578.
 PR 18-FEB-1998; 98US-00025197.
 PR 07-APR-1998; 98US-00056556.
 PR 30-DEC-1998; 98US-00223040.
 XX (REED//) REED S. G.
 PA (SKEI//) SKEIKI Y. A.
 PA (DILL//) DILLON D. C.
 PA (ALDE//) ALDERSON M.
 PA (CAMP//) CAMPOS-NETO A.
 XX
 PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
 XX WPI; 2002-171134/22.
 DR N-PSDB; ABK14139.
 XX
 XX New fusion proteins of Mycobacterium tuberculosis antigens, useful for
 PT diagnosing, treating or preventing M. tuberculosis infection,
 PT particularly as vaccine for treating or preventing tuberculosis.
 XX
 PS Claim 1; Fig 12; 62pp; English.

XX The invention relates to a purified polypeptide which induces an immune
 CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
 CC useful for diagnosing, treating or preventing M. tuberculosis infection,
 CC particularly tuberculosis infection. In particular, the polypeptides are
 CC useful as a vaccine formulation with an adjuvant to afford long-term
 CC protection in animals against the development of tuberculosis. The
 CC protein coding sequence may be used to encode a protein product for use
 CC as an immunogen to induce and/or enhance an immune response to M.
 CC tuberculosis. This sequence represents an M. tuberculosis fusion protein
 CC of the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 599 AA;

Query Match 100.0%; Score 1949; DB 5; Length 599;
 Best Local Similarity 100.0%; Pred. No. 2.1e-142;

Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 60
 Db 9 MVDGALPPEINSARMYAGPGSASLVAAQWDSVASDLFSAASAFQSVVWGLTVGSMIG 68
 QY 61 SSAGLMVAASPYVAVMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120
 Db 69 SSAGLMVAASPYVAVMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 128
 QY 121 LIATNLLGQNTPAIAVNEAEGYEMWAOADAAAFGYAAATATATATLLPFEAPMTSAGG 180
 Db 129 LIATNLLGQNTPAIAVNEAEGYEMWAOADAAAFGYAAATATATATLLPFEAPMTSAGG 188
 QY 181 LLEQAAVEEASDTAAANQLMNVPOALQQAQPTGTTTSSKLGGLWKTVPSPHSPISN 240
 Db 189 LLEQAAVEEASDTAAANQLMNVPOALQQAQPTGTTTSSKLGGLWKTVPSPHSPISN 248
 QY 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 300
 Db 249 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAQAQVTAQAQNGVRAMSSLGSSG 308
 QY 301 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 360
 Db 309 LGGVVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQMLGGLPV 368
 QY 361 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 391
 Db 369 GQMGARAGGGLSGVLRVPPRPVYMPHSPAAG 399

RESULT 15

AAY32068
 ID AAY32068 standard; protein; 600 AA.

AC AAY32068;

DT 17-JAN-2000 (first entry)

DE Mycobacterium tuberculosis antigen fusion protein Mtb61f.

Tuberculosis; antigen; fusion protein; Mtb6.6f; TbH9; DPV; MTI; diagnosis; therapy; vaccine; immunogen.

OS *Mycobacterium tuberculosis*.

PN WO9951748-A2.

PD 14-OCT-1999.

07-APR-1999; 99WO-US007717.

PR 07-APR-1998; 98US-00056556.

FR XX
30-DEC-1998; 9803-00223040.

PA (CORTI-) CORTIXA CORP.
XX

FI
SKEIKY YAW, Alderson M, Campos-Neto A,
XX

DR WPI; 1999-601610/51.
DR N-PSDB: AAZ20203

XX
PT
New fusion protei

XX
XX
PI
PS
tuberculosis.
Claim 1. Fig. 10A-B. 8300. Embodiment

This is
 a
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524

antigen fusion protein, termed Mub611, composed of the antigens TBH9, DP2
and MT1. The fusion protein is expressed in host cells using a vector
carrying a polynucleotide (see AAZ20203) comprising the coding sequences
for the 3 antigens. The invention provides fusion proteins (see AAV32059-
71) containing at least 2 M. tuberculosis antigens. The new fusion for
proteins and polynucleotides encoding them are useful as vaccines for
preventing tuberculosis (claimed), for diagnosis (via in vitro assays or
intradermal skin tests for detection of anti-M. tuberculosis antibodies),
monitoring of disease progression, and treatment of tuberculosis. They
are more effective immunogens than mixtures of the individual protein
components

Sequence 600 AA;

Query Match 100.0%; Score 1949; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 2.1e-142;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MVD FGALPPEIN SARMYAGPGSASLVAAAQMWDSVASDLFSAASFQSVWGLTVGSWIG 60

Db
9 MVDFGALPPEINSARMYAGPGSASLVAAQAQMWDSVASDLFSAASAFQSVWGLTVGSWIG 68

QY 61 SSAGLMVAAASPVAWMSVTAGQAELTAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120

Db 69 SSAGLMVAAASPYYAWMSVTAGQAELTAAQVRVAAAAYETAYGLTVPPPPVTAENRAELMI 128

QY 121 LIATNLLGQNTPAIAVNEAEYCGEMWAQDAAAMFGYAAATATATATLLPFFEEAPEMTSAGG 180

Db 129 LIATNLLGQNTPAIAVNEAEYGEMWAODAAAMFGYAAATATATATLLPFEAPEMTSAGG 188

QY 181 LLEQAAAVEEASDTAAANOLMNNVPOALOOPTOGTTPSSKLGGLWKTVPHRSPISN 240

Db
189 LLEOAAA VEEASDTAAANOLMNNVPOALOO LAOPTOGTTPSSKIGGLWKTTSPHRSPTSN 248

241 MVSMANNHMSMTNSGVSMNTLS SMLKGFA PAAAOAVOTAAONGVRAMSSI.GSSI.GSSG 300

Db 249 MVSMANNHMSMTNSGVSMNTLTSSMLKGFA PAAAAQAVOTAAONGVRAMSSSGSSSG 308

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 03:08:19 ; Search time 34 seconds

(without alignments)
593.699 Million cell updates/sec

Title: US-09-724-685-107

Perfect score: 1949

Sequence: 1 MVDGALPEINSARMYAGP.....SGVLRVPRPYVMHPSPAAG 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/prodata/2/iaa/PCUTUS COMB.pcp.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1949	100.0	391	3	US-08-818-112-107
2	1949	100.0	391	4	US-08-818-111-102
3	1949	100.0	391	4	US-09-056-556-107
4	1949	100.0	391	4	US-09-072-596-102
5	1949	100.0	391	4	US-09-072-567-107
6	1949	100.0	596	4	US-09-287-849-26
7	1949	100.0	600	4	US-09-287-849-22
8	1944	99.7	729	4	US-09-223-040-2
9	1944	99.7	729	4	US-09-287-849-2
10	1652.5	84.8	396	3	US-08-818-112-111
11	1652.5	84.8	396	4	US-08-818-111-106
12	1652.5	84.8	396	4	US-09-056-556-111
13	1652.5	84.8	396	4	US-09-072-596-106
14	1652.5	84.8	396	4	US-09-072-967-111
15	1486.5	76.3	359	3	US-08-818-112-109
16	1486.5	76.3	359	4	US-08-818-111-104
17	1486.5	76.3	359	4	US-09-056-556-109
18	1486.5	76.3	359	4	US-09-072-596-104
19	1486.5	76.3	359	4	US-09-072-967-109
20	1187.5	60.9	358	4	US-09-287-849-8
21	1187	60.9	263	3	US-08-818-112-91
22	1187	60.9	263	4	US-08-818-111-92
23	1187	60.9	263	4	US-09-056-556-91
24	1187	60.9	263	4	US-09-072-596-92
25	1187	60.9	263	4	US-09-072-967-91
26	766.5	39.3	400	4	US-09-073-009-126
27	766.5	39.3	400	4	US-09-073-010-126

28	604	31.0	710	4	US-09-287-849-16
29	604	31.0	856	4	US-09-287-849-12
30	603	30.9	423	4	US-09-073-009-142
31	603	30.9	423	4	US-09-073-010-142
32	424.5	21.8	943	4	US-09-477-135A-131
33	424	21.8	141	4	US-09-073-009-15
34	424	21.8	141	4	US-09-073-010-15
35	381.5	19.6	204	4	US-08-311-731A-57
36	377.5	19.4	208	4	US-08-311-731A-208
37	317	16.3	371	4	US-09-050-739-92
38	314	16.1	368	3	US-08-818-112-114
39	314	16.1	368	4	US-08-818-111-109
40	314	16.1	368	4	US-09-056-556-114
41	314	16.1	368	4	US-09-072-596-109
42	314	16.1	368	4	US-09-072-967-114
43	261	13.4	1271	1	US-08-095-734-2
44	261	13.4	1271	2	US-08-444-623-2
45	261	13.4	1271	3	US-08-471-869-2

ALIGNMENTS

RESULT 1
US-08-818-112-107
; Sequence 107, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 822-4900
; TELEFAX: (206) 822-6031
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-818-112-107

Query Match 100.0%; Score 1949; DB 3; Length 391;
Best Local Similarity 100.0%; Pred. No. 2.9e-154;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVDGALPEINSAARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60
DB 1 MVDGALPEINSAARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60
QY 61 SSAGLMVAASAPYVAMSVTAAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLMVAASAPYVAMSVTAAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
QY 121 LIATNLLGONTPTAVNEAEYGEWMAQDAAMFGYAAATATATATLTPFEAPEMTSAGG 180
DB 121 LIATNLLGONTPTAVNEAEYGEWMAQDAAMFGYAAATATATATLTPFEAPEMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240
DB 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240
QY 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300
DB 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300
QY 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQWMLGGLPV 360
DB 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQWMLGGLPV 360
QY 361 QMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391
DB 361 QMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391

```

```

RESULT 2
US-08-818-111-102
; Sequence 102, Application US/08818111
; Patent No. 6338852
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,111
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C6
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-818-111-102

```

```

Query Match 100.0%; Score 1949; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 2.9e-154;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDGALPEINSAARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60
DB 1 MVDGALPEINSAARMYAGPGSASLVAAQWMDSVASDLFSAASAFOSVWGLTVGSMIG 60
QY 61 SSAGLMVAASAPYVAMSVTAAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLMVAASAPYVAMSVTAAQAEELTAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
QY 121 LIATNLLGONTPTAVNEAEYGEWMAQDAAMFGYAAATATATATLTPFEAPEMTSAGG 180
DB 121 LIATNLLGONTPTAVNEAEYGEWMAQDAAMFGYAAATATATATLTPFEAPEMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240
DB 181 LLEQAAVEEASDTAAANQLMNNVPQALQQLAQTQGTTPSSKLGGLWKTVPSPHRSPI 240
QY 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300
DB 241 MYSMANHMSMTNSGVSMNTLSSMLKGFAPAPAAAQAVQTAQNGVRAMSSLSGSSG 300
QY 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQWMLGGLPV 360
DB 301 LGGGVAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQWMLGGLPV 360
QY 361 QMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391
DB 361 QMGARAGGGLSGVLRVPPRPYVMPHSPAAG 391

```

```

RESULT 3
US-09-056-556-107
; Sequence 107, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
TREATME

```


US-09-072-967-107
; Sequence 107, Application US/09072967
; Patent No. 6592877
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendricksen, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND
; METHODS FOR IMMUNOTHERAPY
; OF TUBERCULOSIS
; TITLE OF INVENTION: AND DIAGNOSIS OF
; TUBERCULOSIS
; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967


```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-09-287-849-22

Query Match 100.0%; Score 1949; DB 4; Length 600;
Best Local Similarity 100.0%; Pred. No. 5.2e-154; Indels 0; Gaps 0;
Matches 391; Conservative 0; Mismatches 0;

QY 1 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
Db 9 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 68
QY 61 SSAGLMVAASAPYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPPIAENRAELMI 120
Db 69 SSAGLMVAASAPYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPPIAENRAELMI 128
QY 121 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATATLLPPEEAPMTSAGG 180
Db 129 LIATNLLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 188
QY 181 LLEQAAAVEASDTAAANQLMNNVPQALQOLAQPTQGTTPSSKLGGLWKTYSPPHRSPI 240
Db 249 LLEQAAAVEASDTAAANQLMNNVPQALQOLAQPTQGTTPSSKLGGLWKTYSPPHRSPI 248
QY 241 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLSGSSG 300
Db 249 MVSMAHNMHMTNSGVSMNTLSSMLKGFAPAAAAQAVQTAQAQNGVRAMSSLSGSSG 308
QY 301 LGGVVAANLGRAASVGSLSVQFAWAAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 360
Db 309 LGGVVAANLGRAASVGSLSVQFAWAAANQAVTPAARALPLTSLTSAABERGPGQMLGGLPV 368
QY 361 GQMGARAGGSLGVLRVPPRPYVMPHSPAAG 391
Db 369 GQMGARAGGSLGVLRVPPRPYVMPHSPAAG 399

RESULT 8
US-09-223-040-2
; Sequence 2, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009010US
; CURRENT FILING DATE: US/09/223,040
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-09-223-040-2

Query Match 99.7%; Score 1944; DB 4; Length 729;
Best Local Similarity 99.7%; Pred. No. 1.8e-153;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
Db 142 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 201
QY 61 SSAGLMVAASAPYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPPIAENRAELMI 120
Db 142 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 201

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-09-287-849-2

Query Match 99.7%; Score 1944; DB 4; Length 729;
Best Local Similarity 99.7%; Pred. No. 1.8e-153;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 60
Db 142 MVDFFGALPPEINSARMYAGPGSASLVAAQAQWMDSVASDLFSAASAFQSVVWGLTVGSGWIG 201
QY 61 SSAGLMVAASAPYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPPIAENRAELMI 120
Db 202 SSAGLMVAASAPYVAMSVTAGQAEITAAQVRVAAAAAYETAYGLTVPPPIAENRAELMI 261
```


Query Match 84.8%; Score 1652.5; DB 4; Length 396;
Best Local Similarity 84.9%; Pred. No. 1.4e-129;
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60
DB 1 VVDFGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60
QY 61 SSAGLVAAASPYYAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLVAAASPYYAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
DB 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQQAQPTGTTSSKLGGLWKTVPSPHRSPTSN 240
DB 181 LLEQAAVEEADITAAANQLMNNVPQALQQAQPTGTTSSKLGGLWKTVPSPHRSPTSN 240
QY 241 MVSMAHNSMTNSGVSMNTLSSMLKGFAPAAAQAVOTAAQNGVRAMSS---LGSSL 296
DB 241 IVSMLNHNVSMTNSGVSMNTLSSMLKGFAPAAAQAVOTAAQNGVRAMSS---LGSSL 296
QY 297 GSSGLGGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQOMLG 356
DB 300 GSSGLGAGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQOMLG 359
QY 357 GLPVGQMGARAG--GGLSGVLVRPPRYVMPHSPAAG 391
DB 360 GLPLGQLTNSGGGFGVSNALRMPPRAYVMPRVPAAG 396

RESULT 12

US-09-056-556-111
; Sequence 111, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,556
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-056-556-111

TREATM

Query Match 84.8%; Score 1652.5; DB 4; Length 396;
Best Local Similarity 84.9%; Pred. No. 1.4e-129;
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60
DB 1 VVDFGALPPEINSARMYAGPGSASLVAAAQWMDSVASDLFSAASAFQSVVWGLTVGSMIG 60
QY 61 SSAGLVAAASPYYAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
DB 61 SSAGLVAAASPYYAMWSVTAGQAEITAAQVRVAAAAYETAYGLTVPPVIAENRAELMI 120
QY 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
DB 121 LIATNLGQNTPAIAVNEAEYGEWMAQDAAMFGYAAATATATATLLPPEEAPMTSAGG 180
QY 181 LLEQAAVEEASDTAAANQLMNNVPQALQQAQPTGTTSSKLGGLWKTVPSPHRSPTSN 240
DB 181 LLEQAAVEEADITAAANQLMNNVPQALQQAQPTGTTSSKLGGLWKTVPSPHRSPTSN 240
QY 241 MVSMAHNSMTNSGVSMNTLSSMLKGFAPAAAQAVOTAAQNGVRAMSS---LGSSL 296
DB 241 IVSMLNHNVSMTNSGVSMNTLSSMLKGFAPAAAQAVOTAAQNGVRAMSS---LGSSL 296
QY 297 GSSGLGGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQOMLG 356
DB 300 GSSGLGAGVAAANLGRAASVGSLSVPOAWAANAQVTPAARALPLTSLTSAERGPQOMLG 359
QY 357 GLPVGQMGARAG--GGLSGVLVRPPRYVMPHSPAAG 391
DB 360 GLPLGQLTNSGGGFGVSNALRMPPRAYVMPRVPAAG 396

RESULT 13

US-09-072-596-106
; Sequence 106, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-072-596-106

Query Match 84.8%; Score 1652.5; DB 4; Length 396;
Best Local Similarity 84.9%; Pred. No. 1.4e-129;
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

```

QY 1 MVDFGALPPEINSAARMYAGPGSASLVAAAKWDSVASDLFSAASAFQSVVWGLTVGWSIG 60
DB 1 VVDFGALPPEINSAARMYAGPGSASLVAAAKWDSVASDLFSAASAFQSVVWGLTVGWSIG 60

QY 61 SSAGLMVAAASPYVAAWMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120
DB 61 SSAGLMVAAASPYVAAWMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120

QY 121 LIATNLLQONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATATLPPFEAPEMTSAGG 180
DB 121 LIATNLLQONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTQGTTPSSKLGGLWKTVPSPHSPI 240
DB 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTQGTTPSSKLGGLWKTVPSPHSPI 240

QY 241 MVSMAHNMHMTNSGVSMNTLTSSMLKGFAPAAAQAVQTAAGVRAVSS----LGSSL 296
DB 241 MVSMAHNMHMTNSGVSMNTLTSSMLKGFAPAAAQAVQTAAGVRAVSS----LGSSL 296

QY 297 GSSGLGGVAAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERPGOMLG 356
DB 297 GSSGLGGVAAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERPGOMLG 356

QY 357 GLPVGQMGARAG--GGLSGVLRVPPRYVMPHSPAAG 391
DB 360 GLPLGQLTNSGGFGGVSNALRMPPRAYVMPRVPAAG 396

```

RESULT 14

US-09-072-967-111
Sequence 111, Application US/09072967
Patent No. 5592877

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 355

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/072,967

FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Maxi, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-072-967-111

```

Query Match 84.8%; Score 1652.5; DB 4; Length 396;  
Best Local Similarity 84.9%; Pred. No. 1.4e-129;  
Matches 337; Conservative 19; Mismatches 34; Indels 7; Gaps 3;

QY 1 MVDFGALPPEINSAARMYAGPGSASLVAAAKWDSVASDLFSAASAFQSVVWGLTVGWSIG 60
DB 1 VVDFGALPPEINSAARMYAGPGSASLVAAAKWDSVASDLFSAASAFQSVVWGLTVGWSIG 60

QY 61 SSAGLMVAAASPYVAAWMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120
DB 61 SSAGLMVAAASPYVAAWMSVTAGQAEELTAQVRVAAAAYETAYGLTVPPPIAENRAELMI 120

QY 121 LIATNLLQONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180
DB 121 LIATNLLQONTPTAIVNEAEYGEWMAQDAAMFGYAAATATATATLPPFEAPEMTSAGG 180

QY 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTQGTTPSSKLGGLWKTVPSPHSPI 240
DB 181 LLEQAAVEEASDTAAANQLMNNVPOALQQLAQTQGTTPSSKLGGLWKTVPSPHSPI 240

QY 241 MVSMAHNMHMTNSGVSMNTLTSSMLKGFAPAAAQAVQTAAGVRAVSS----LGSSL 296
DB 241 MVSMAHNMHMTNSGVSMNTLTSSMLKGFAPAAAQAVQTAAGVRAVSS----LGSSL 296

QY 297 GSSGLGGVAAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERPGOMLG 356
DB 297 GSSGLGGVAAANLGRAASVGSLSVPQAWAANQAVTPAARALPLTSLTSAERPGOMLG 356

QY 357 GLPVGQMGARAG--GGLSGVLRVPPRYVMPHSPAAG 391
DB 360 GLPLGQLTNSGGFGGVSNALRMPPRAYVMPRVPAAG 396

```

RESULT 15

US-08-818-112-109
Sequence 109, Application US/08818112
Patent No. 6290969

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

Search completed: August 25, 2004, 03:14:26
Job time : 36 secs

This Page Blank (uspto)